

Universidad Autónoma de Madrid
Departamento de Bioquímica

microRNAs en el desarrollo temprano de ratón:
análisis de la falta de función de Dicer en el linaje del trofoblasto.

Bárbara Pernaute Lomba
Madrid, 2009

Departamento de Bioquímica
Facultad de Medicina
Universidad Autónoma de Madrid

microRNAs en el desarrollo temprano de ratón:
análisis de la falta de función de Dicer en el linaje del trofoblasto.

Bárbara Pernaute Lomba. Licenciada en Biología.

Director de la tesis: Miguel Manzanares Fourcade.

Lugar de Realización: Instituto de Investigaciones Biomédicas “Alberto Sols” y
Fundación Centro Nacional de Investigaciones Cardiovasculares Carlos III

Madrid



MINISTERIO
DE CIENCIA
E INNOVACIÓN



cnic

Yo, MIGUEL MANZANARES FOURCADE, Investigador del Centro Nacional de Investigaciones Cardiovasculares (CNIC) y Profesor Honorario del Departamento de Bioquímica de la Universidad Autónoma de Madrid,

certifico que

BÁRBARA PERNAUTE LOMBA, licenciada en Biología por la Universidad Complutense de Madrid, ha realizado bajo mi dirección y supervisión el trabajo de investigación titulado

**“microRNAs en el desarrollo temprano de ratón:
análisis de la falta de función de Dicer en el linaje del trofoblasto”**

Considero, además, que el mencionado trabajo reúne la originalidad y calidad científica requeridas para poder ser presentado para optar al grado de Doctor por la Universidad Autónoma de Madrid.

Y para que conste a todos los efectos, firmo el presente certificado,

En Madrid, a 25 de Mayo de 2009

Miguel Manzanares
Department of Cardiovascular Developmental Biology
Centro Nacional de Investigaciones Cardiovasculares-CNIC
Melchor Fernandez Almagro, 3
28029 Madrid, Spain

tel: (34) 91 453 13 11
fax: (34) 91 453 13 04
mmanzanares@cnic.es
www.cnic.es/cdbd

A mis padres

Resumen

Los procesos que tienen lugar durante el desarrollo embrionario temprano son fundamentales para el correcto establecimiento del patrón de cualquier organismo. La aparición de un desarrollo completamente intrauterino en mamíferos conlleva la formación de una estructura extraembrionaria altamente especializada y necesaria no sólo para el soporte del embrión sino también para su correcta formación en las etapas tempranas: el trofoblasto, precursor de la placenta. Los eventos que tienen lugar durante el desarrollo de las estructuras embrionarias y extraembrionarias están finamente regulados a nivel transcripcional y post-transcripcional. Entre los mecanismos de regulación post-transcripcional se encuentra la represión mediada por microRNAs (miRNAs), RNAs de pequeño tamaño que se unen de forma complementaria a mRNAs diana provocando su degradación o impidiendo su traducción, y que se encuentran implicados en multitud de procesos entre los que se incluye el desarrollo.

En este trabajo hemos abordado el estudio de la función de los miRNAs durante las primeras etapas del desarrollo embrionario de ratón, centrando nuestra atención sobre el linaje extraembrionario del trofoblasto. Para ello hemos analizado el trofoblasto de embriones mutantes completos para la proteína *Dicer*, esencial en la maduración de estos represores y cuya delección implica la ausencia de miRNAs maduros. Asimismo, hemos derivado dos líneas de células madre de trofoblasto (células TS) condicionales para *Dicer*, que nos han permitido deleccionar el gen en cultivo y analizar el fenotipo de este tipo celular durante los primeros momentos tras la retirada de los miRNAs.

El análisis del fenotipo de falta de función de *Dicer* en el trofoblasto embrionario y en células TS, en combinación con la obtención por primera vez de los perfiles de expresión de miRNAs en embriones en post-implantación y en células madre de trofoblasto, nos ha servido para describir una función de los miRNAs en el mantenimiento de la población de células TS *in vivo* e *in vitro*. La delección de *Dicer* provoca una drástica bajada de la proliferación de las células TS posiblemente debida a la ausencia de miRNAs implicados en el control del ciclo celular mediante un mecanismo similar al ya descrito para células madre embrionarias (ES). Por otro lado y al contrario de lo que sucede en células ES, la delección de *Dicer* provoca la diferenciación de las células TS en los distintos tipos celulares del linaje del trofoblasto, entre los que se encuentran las células gigantes del trofoblasto (TGCs). La diferenciación hacia TGCs implica la endo-reduplicación programada del DNA, un fenómeno que raramente se da en las células de mamíferos y que de acuerdo a nuestros resultados podría estar controlado al menos en parte por miRNAs cuya expresión es predominante en el linaje del trofoblasto.

Por último, los resultados obtenidos en relación al fenotipo de falta de función de *Dicer* en embriones en post-implantación muestran que si bien los miRNAs son de vital importancia para el correcto desarrollo embrionario temprano, la función individual de cada uno de ellos va encaminada a la modulación de la expresión génica, confiriendo un grado extra de solidez a redes regulatorias en las que intervienen otros mecanismos como la regulación transcripcional.

Abstract

The processes that take place during early embryo development are critical for pattern establishment in all organisms. Intrauterine development in mammals requires the formation of a highly specialized extraembryonic structure: the trophoblast that will give rise to the placenta. The events that take place during the development of embryonic and extraembryonic structures are precisely regulated at both transcriptional and post-transcriptional levels. Among the post-transcriptional regulation mechanisms is the repression of gene expression mediated by microRNAs (miRNAs). miRNAs are small non-coding RNAs that bind in a complementary way to their target mRNAs leading to their degradation or impairing their translation. They have been found to be involved in many different processes including embryo development.

In this work we have conducted a study of miRNA function during the early stages of mouse embryo development, focusing our attention on the extraembryonic lineage of the trophoblast. In order to do that we have analyzed the trophoblast of embryos in which the protein Dicer has been knocked out. Dicer is essential for miRNA maturation and its deletion results in the absence of mature miRNAs. Additionally we have derived two trophoblast stem cells (TS cells) lines in which *Dicer* can be conditionally deleted. These have allowed us to analyse the effects of miRNA withdrawal on this cell type in culture.

The analysis of the Dicer lack of function phenotype in the trophoblast and in TS cells combined with the characterization of miRNA expression profiles in early post-implantation embryos and in trophoblast stem cells, has allowed us to describe a function for miRNAs in maintaining the TS cells population both *in vivo* and *in vitro*. *Dicer* deletion leads to a severe decrease in the proliferation of TS cells. This is possibly due to the loss of miRNAs involved in cell cycle control by a mechanism similar to the one that has been proposed to account for the proliferation defects that are seen in *Dicer* mutant embryonic stem (ES) cells. Moreover and contrary to what happens in ES cells, *Dicer* deletion leads to the differentiation of TS cells into all the cell subtypes of the trophoblast lineage, including trophoblast giant cells (TGCs). Differentiation into TGCs requires endo-reduplication of the DNA, a phenomenon that rarely occurs in mammalian cells and that according to our results could be in part controlled by miRNAs predominantly expressed in the trophoblast lineage.

Finally the results obtained from the phenotype analysis of post-implantation *Dicer* null embryos show that although miRNAs are crucial for the correct development of the early embryo, the individual function of each of them seems to be to add an extra level of robustness to regulatory networks in which mechanisms of regulation at a transcriptional level are involved.

Indice

Indice

Abreviaturas	5
Introducción	9
Desarrollo embrionario temprano	11
1 Desarrollo en pre-implantación: especificación de los primeros linajes	11
2 Desarrollo en post-implantación: etapas previas a la gastrulación	13
3 Expansión del trofoblasto y formación de la placenta	16
4 Células madre derivadas del blastocisto	16
4.1 Células madre de trofoblasto	16
Regulación post-transcripcional mediada por miRNAs	18
1 Organización genómica de los miRNAs	19
2 Biogénesis de los miRNAs	19
2.1 Transcripción	20
2.2 Procesamiento del pri-miRNA	21
2.3 Procesamiento del pre-miRNA	21
2.4 Represión mediada por miRNAs	22
Función de los miRNAs en el desarrollo	23
1 Análisis de expresión de los miRNAs y búsqueda de mRNAs diana	23
1.1 Análisis de expresión	23
1.2 Búsqueda de mRNAs diana	24
2 Implicación de los miRNAs en el desarrollo embrionario	24
3 Mutantes de <i>Dicer</i> : modelo para el estudio de la función de los miRNAs en el desarrollo del ratón	27
3.1 Función de los miRNAs en el desarrollo embrionario temprano <i>in vivo</i>	28
3.2 Función de los miRNAs <i>in vitro</i> : análisis en células madre embrionarias	29
Objetivos	31
Materiales y Métodos	35
Análisis <i>in vivo</i> del fenotipo de falta de función de Dicer en el trofoblasto	37
1 Líneas de ratón y genotipaje	37
2 Hibridación <i>in situ</i> en embriones enteros	37
3 Análisis de proliferación con fosfo-histona 3 (PH3)	38

Análisis <i>in vitro</i> : fenotipo de las células TS- <i>Dicer</i> ^{-/-}	39
4 Derivación de las células TS- <i>Dicer</i> ^{fx/fx}	39
4.1 Obtención de las líneas TSL1- <i>Dicer</i> ^{fx/fx} y TSL4- <i>Dicer</i> ^{fx/fx}	39
4.2 Caracterización de las líneas TSL1- <i>Dicer</i> ^{fx/fx} y TSL4- <i>Dicer</i> ^{fx/fx}	40
5 Deleción de <i>Dicer</i> <i>in vitro</i> : sistema de infección con adenovirus	41
6 Análisis de las células TS- <i>Dicer</i> ^{-/-}	43
6.1 Recuento del número de células y extracción de RNA	43
6.2 PCR cuantitativa	43
7 Variación en mRNA en células TS- <i>Dicer</i> ^{-/-} : array de mRNA	44
7.1 Análisis de Gene Ontology (GO)	45
Expresión de miRNAs en células TS y embrión temprano	45
8 Análisis de expresión de miRNAs: array por PCR Cuantitativa	45
8.1 Análisis de datos del array de miRNAs	46
8.2 Patrones de expresión de miRNAs	47
miRNAs implicados en la función de las células TS	48
9 miRNAs asociados a la variación génica de las células TS- <i>Dicer</i> ^{-/-}	48
Resultados	51
Análisis <i>in vivo</i> del fenotipo de falta de función de Dicer en el trofoblasto	53
1 Los embriones <i>Dicer</i> ^{-/-} son morfológicamente anormales y mueren en gastrulación	53
2 Defectos en el trofoblasto de embriones <i>Dicer</i> ^{-/-}	54
2.1 Falta de expresión de genes de multipotencia	54
2.2 El fenotipo del trofoblasto es independiente del epiblasto	56
2.3 El trofoblasto de los embriones <i>Dicer</i> ^{-/-} presenta defectos de proliferación	58
Análisis <i>in vitro</i> : fenotipo de las células TS- <i>Dicer</i> ^{-/-}	59
3 Las células TS- <i>Dicer</i> ^{-/-} tienen defectos en el mantenimiento de la multipotencia y la proliferación	59
4 Cambios en RNA mensajero en células TS tras la deleción de <i>Dicer</i>	62
4.1 Rutas de señalización implicadas en la diferenciación	65
4.1.1 Señalización por Fgf y TGFbeta	65
4.1.2 Otros factores implicados en diferenciación	66
4.1.3 Regulación por Cdkn1c	67
4.2 Ruta de señalización implicada en la bajada de proliferación	67
Expresión de miRNAs en células TS y embrión temprano	68
5 Perfil de expresión de miRNAs en células TS y en embrión temprano	69
6 Perfil de expresión en relación a la organización genómica	71

6.1 Patrones de expresión de miRNAs restringidos	72
miRNAs implicados en la función de células TS	75
7 miRNAs asociados a la variación en expresión génica de las células TS- <i>Dicer</i> ^{-/-}	76
8 miRNAs relacionados con el fenotipo de las células TS- <i>Dicer</i> ^{-/-}	78
Discusión	81
1 Los miRNAs se requieren para mantener las células TS en el embrión	83
2 Papel de los miRNAs en la proliferación de las células TS	84
3 Papel de los miRNAs en el mantenimiento de la multipotencia de las células TS	87
4 Control del ciclo celular de las células TS por miRNAs	89
5 Diferencias en el control de la diferenciación por miRNAs en células ES y TS	91
6 Función de los miRNAs en el desarrollo temprano del embrión de ratón	92
6.1 miRNAs como moduladores de la expresión génica durante el desarrollo	95
Conclusiones	97
Bibliografía	101
Apéndices	113

Abreviaturas

AVE	Endodermo Visceral Anterior
EP	Endodermo Primitivo
EPC	Cono Ectoplacentario
Epi	Epiblasto
Células ES	Células Madre Embrionarias
Células TS	Células Madre de Trofoblasto
Células XEN	Células madre de Endodermo Extraembrionario
CM	Células madre derivadas del blastocisto
EB	Cuerpo embrionario
ExE	Ectodermo Extraembrionario
exVE	Endodermo Visceral Extraembrionario
Fx	Alelo condicional para <i>Dicer</i>
GO	Ontología génica
ICM	Masa Celular Interna
LNA	Acido nucleico modificado utilizado para la detección de expresión de microARNs
LP	Línea Primitiva
MEFs	Fibroblastos embrionarios de ratón
miRNAs	microARNs
TE	Trofoectodermo
VE	Endodermo Visceral
Wt	Fenotipo silvestre
UTR	Región del ARN mensajero no traducida
TGCs	Células gigantes del trofoblasto
RISC	Complejo de silenciamiento inducido por ARN de interferencia
siRNAs	ARNs pequeños de interferencia

Introducción

El desarrollo embrionario es el proceso mediante el cual se especifican todos los linajes que conforman un organismo a partir de una única célula. Esta especificación de linajes ocurre bajo un perfecto control espacio-temporal de procesos como diferenciación, proliferación y muerte celular, que han de estar finamente regulados de manera que se siga un patrón correcto de formación del embrión. Durante las etapas más tempranas del desarrollo embrionario tienen lugar decisiones tan importantes como la especificación de los primeros linajes o el establecimiento de los ejes del embrión. En los mamíferos el embrión se caracteriza por presentar un desarrollo intrauterino durante el cual el aporte nutricional depende exclusivamente de la madre. Por ello no es de extrañar que las primeras decisiones de linaje que tienen lugar vayan encaminadas a la formación de estructuras de soporte del embrión: éstos son el trofoblasto que dará lugar a la placenta, y el endodermo primitivo que contribuirá a la formación del saco vitelino (Rossant, 2004; Rossant and Tam, 2009). En el ratón los tejidos extraembrionarios, además de constituir estructuras de soporte imprescindibles, van a jugar un papel fundamental en el establecimiento del patrón del embrión temprano, ya que linajes embrionario y extraembrionario se comunican entre sí y contribuyen mutuamente a su desarrollo (Rossant and Tam, 2009). Por este motivo no es posible entender los procesos que tienen lugar durante el desarrollo temprano del embrión fuera del contexto de los que suceden en el desarrollo de las estructuras extraembrionarias.

El presente trabajo forma parte de un proyecto cuyo objetivo principal es el análisis de la importancia de la regulación post-transcripcional mediada por microRNAs (miRNAs) en el desarrollo de los tres linajes que constituyen el embrión temprano de ratón. Dicho proyecto se ha realizado en colaboración con el laboratorio del Dr. Tristán Rodríguez, donde se han llevado a cabo los análisis referentes a la función de los miRNAs en el desarrollo del epiblasto y el endodermo visceral. El trabajo aquí presentado se centra por contra en el estudio de la regulación que tiene lugar en el trofoblasto. Durante los siguientes apartados describiremos los procesos fundamentales que tienen lugar durante el desarrollo temprano del embrión de ratón haciendo especial hincapié en la formación del trofoblasto y sus derivados; asimismo haremos una revisión de los conocimientos que hasta el momento se tienen acerca de la función de los miRNAs, especialmente en el contexto del desarrollo embrionario.

DESARROLLO EMBRIONARIO TEMPRANO

1 Desarrollo en pre-implantación: especificación de los primeros linajes

Durante los cuatro primeros días tras la fecundación el embrión de ratón se desarrolla sin establecer contacto con el útero materno, pasando por lo que se denominan etapas de pre-implantación.

El oocito fecundado de ratón sufre una primera división simétrica que da lugar al embrión de dos células. Este momento es destacable porque aquí tiene lugar la activación del genoma zigótico acompañada de la degradación del mRNA materno presente inicialmente en el oocito (Schultz, 2002). Las dos siguientes divisiones, también simétricas, dan lugar a la formación del embrión de cuatro y ocho células respectivamente. A los 2.5 días de desarrollo el embrión de ocho células se compacta y las células adquieren una morfología de epitelio polarizado que se va a mantener únicamente en las células más externas durante las sucesivas divisiones para formar el embrión de 16 y 32 células, conocido como mórula compacta. Esta distinción entre células externas polarizadas e internas no polarizadas va a dar lugar a la primera decisión de linaje que tiene lugar en el desarrollo del embrión de ratón y que supone la especificación del trofoectodermo que dará lugar al trofoblasto. Así, a los 3.5 días de desarrollo la mórula compacta se expande para formar una cavidad conocida como blastocelo y una masa de células internas (la masa celular interna o ICM, del inglés Inner Cell Mass) precursoras del embrión, ambos rodeados por las células externas que forman ahora una capa epitelial, el trofoectodermo a partir del cual se desarrolla el trofoblasto. El embrión en esta etapa del desarrollo se conoce como blastocisto, y está formado únicamente por los linajes de la ICM y el trofoectodermo (Kunath et al., 2004). Mientras que el trofoectodermo supone un linaje exclusivamente extraembrionario, las células de la masa celular interna van a sufrir otra segregación de linajes durante la etapa de blastocisto tardío, por la cual se distinguen las células que formarán el endodermo primitivo de aquellas que únicamente darán lugar a linajes embrionarios y que constituyen el epiblasto (Rossant and Tam, 2009). De este modo, en el blastocisto tardío encontramos dos linajes extraembrionarios, trofoectodermo y endodermo primitivo, diferenciados del epiblasto que va a formar en exclusiva todas las estructuras del embrión.

La segregación de estos linajes está finamente regulada por una red génica en la que intervienen factores de transcripción expresados específicamente en cada tipo celular y que actúan potenciando un determinado destino celular a la vez que reprimen el contrario. El mecanismo exacto por el que se desencadena la especificación del linaje del trofoectodermo en las células más externas de la mórula compacta es aún desconocido. Se ha implicado recientemente al factor de transcripción *Tead4* como primer responsable conocido de la especificación del trofoectodermo a través de su interacción con proteínas de la vía Hippo, que hacen que se localice de forma nuclear únicamente en las células más externas de la mórula compacta (Nishioka et al., 2009; Yagi et al., 2007). Sin embargo, el primer gen con expresión restringida al linaje del trofoectodermo descrito hasta el momento es el factor de transcripción *Cdx2* (Strumpf et al., 2005). En la formación inicial del blastocisto *Cdx2* establece una relación de mutua represión con el factor de transcripción Oct4, codificado por el gen *Pou5f1* y cuya expresión se restringe a células de la masa celular interna. Por otro lado, *Cdx2* dirige la expresión en el trofoectodermo del factor de transcripción *Eomes*, esencial también para el desarrollo del trofoecodermo (Niwa et al., 2005; Russ et al., 2000). En la segunda decisión de

linaje, por la cual se van a formar el epiblasto y en endodermo primitivo, intervienen de nuevo dos factores de transcripción específicos de tipo celular, que son en este caso *Nanog* en el epiblasto y *Gata6* en el endodermo primitivo. Del mismo modo que *Oct4* actúa en el mantenimiento del destino celular embrionario frente al de trofoectodermo mediante la represión de *Cdx2*, *Nanog* lo mantiene frente al de endodermo primitivo mediante la represión de *Gata6* (Rossant and Tam, 2009). Así pues, en el blastocisto *Cdx2* y *Eomes* se encargan del mantenimiento de las células del trofoblasto mientras que *Pou5f1* (*Oct4*) y *Nanog* se ocupan de preservar el destino embrionario.

2 Desarrollo en post-implantación: etapas previas a la gastrulación

En la figura 1 se encuentra representado un esquema de los distintos linajes especificados en las etapas comprendidas entre el estadio de blastocisto y el fin de la gastrulación a las 7.5 días de desarrollo. A partir del día 4 del desarrollo el embrión tienen una conformación de blastocisto tardío y es en este momento cuando tiene lugar la implantación en el útero, que se da a través de la región de trofoectodermo que rodea a la ICM, denominada “trofoectodermo polar”. Las células que conforman la región opuesta al trofoectodermo polar, llamada “trofoectodermo mural”, endo-reduplican su DNA y se diferencian en las llamadas células gigantes del trofoblasto (TGCs, del inglés *Trophoblast Giant Cells*) (Kunath et al., 2004).

En este momento, las células del epiblasto proliferan y se desplazan hacia el polo distal del blastocisto reduciendo la cavidad del blastocele y empujando al endodermo primitivo, que quedará formando una capa que rodea tanto al epiblasto como al trofoblasto. Una parte de las células que componen el endodermo primitivo se diferenciará en el endodermo parietal, que se une a las TGCs generadas a partir del trofoectodermo mural para formar una capa que dará lugar más adelante al saco vitelino, mientras que las células que permanecen en contacto con el epiblasto y trofoblasto conforman el endodermo visceral embrionario y extraembrionario respectivamente. Al mismo tiempo que esto sucede, las células del trofoectodermo polar proliferan para formar el trofoblasto precursor de todos los linajes de la placenta (Rossant, 2004).

En el trofoblasto del embrión de estadio E5.5 y E6.5 encontramos varias poblaciones celulares que difieren en el grado de diferenciación que presentan. Así, próximo al epiblasto se localiza el tejido del ectodermo extraembrionario (ExE, del inglés *Extraembryonic Ectoderm*) en el que encontramos, entre otras, la población de células TS (del inglés, *Trophoblast Stem cells*), que se mantiene como reservorio de precursores indiferenciados hasta al menos estadio E7.5 (Tanaka et al., 1998). Aunque no se sabe con precisión en dónde se localiza esta población de células dentro del ExE, la expresión de marcadores de multipotencia propios de ellas como son *Cdx2* y *Esrrb* parece indicar que se encuentran en la zona más próxima al epiblasto. La porción del trofoblasto más diferenciada es la que mantiene un contacto más directo con la decidua materna: el cono ectoplacentario (EPC, del inglés, *Ectoplacental Cone*). Las células del EPC

más próximas a la pared del útero endo-reduplican su DNA diferenciándose así en células gigantes de trofoblasto (Rossant and Cross, 2001).

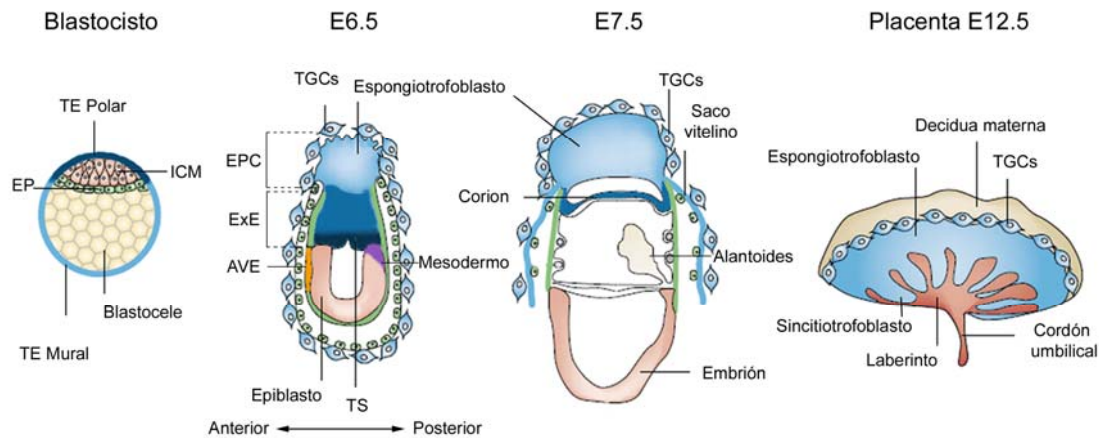


Figura 1. Desarrollo temprano del embrión y la placenta de ratón. El embrión de ratón a los 4 días de desarrollo (blastocisto tardío) ha diferenciado dos linajes extraembrionarios (trofoectodermo, TE; y endodermo primitivo, EP) y uno embrionario (masa celular interna, ICM). Las células del trofoectodermo polar (TE polar) que rodean a la masa celular interna continúan proliferando y forman en E6.5 el trofoblasto, mientras que las del trofoectodermo mural (TE mural) se diferencian en células gigantes de trofoblasto (TGCs). El trofoblasto en E6.5 está compuesto por el ectodermo extraembrionario (ExE) y el cono ectoplacentario (EPC). Dentro del ectodermo extraembrionario encontramos una población de precursores indiferenciados del trofoblasto: las células madre de trofoblasto (células TS). En este estadio y debido a la migración del endodermo visceral anterior (AVE), se establece en el epiblasto el eje antero-posterior, especificándose el mesodermo en la región posterior. En estadio E7.5 el ExE ha dado lugar al ectodermo del corion y el EPC al espongiotrofoblasto, mientras que de la parte posterior embrionaria emerge el alantoides. La fusión del corion y el alantoides dará lugar en estadios posteriores a la región del laberinto, formado por la red vascular que permite el intercambio entre la madre y el feto y que se encuentra rodeada por las células del sincitiotrofoblasto. Más externamente queda el espongiotrofoblasto, rodeado a su vez por las TGCs que establecen contacto con la decidua materna. Modificado de (Rossant and Cross, 2001)

En las fases del desarrollo comprendidas entre E5.5 y E6.5 la comunicación entre los dos tejidos extraembrionarios (endodermo visceral y trofoblasto) y el tejido embrionario (epiblasto) resulta de vital importancia para el desarrollo de los tres linajes. Existen factores clave expresados en las células TS que son imprescindibles para su mantenimiento, éstos son principalmente *Cdx2*, *Eomes* y el receptor de estrógenos *Esrrb*. La falta de expresión de cualquiera de estos factores va a tener como consecuencia la pérdida de la multipotencia y capacidad de auto-renovación de estas células, resultando en el impedimento del desarrollo del trofoblasto y la muerte del embrión (Luo et al., 1997; Russ et al., 2000; Strumpf et al., 2005). El epiblasto produce factores que son fundamentales en el mantenimiento de la población de células TS dentro del ectodermo extraembrionario (Fig. 2). Estos son el factor de crecimiento Fgf4 y el miembro de la familia TGFbeta Nodal. Fgf4 producido en el epiblasto se une a su receptor Fgfr2, expresado en el ectodermo extraembrionario, y activa una vía de señalización resultante en la fosforilación y activación de Erk1/2. La señalización a través de esta vía actúa promoviendo la expresión de *Cdx2*, *Eomes* y *Esrrb*, y restringe a su vez la expresión del promotor de diferenciación *Asc2* al territorio del EPC (Rossant and Cross, 2001; Tanaka et al.,

1998). *Nodal* por su parte se expresa también en el epiblasto del embrión temprano en post-implantación. La proteína Nodal es procesada en el ExE por las convertasas Spc1 y Spc4, expresadas específicamente en este tejido. Nodal procesado ejerce a continuación su acción tanto en el epiblasto como en el ExE. En el primero activa vías de señalización implicadas en el establecimiento del patrón embrionario, mientras que en el ExE es necesario para el mantenimiento de la expresión de *Cdx2*, *Eomes* y *Esrrb* (Guzman-Ayala et al., 2004). La importancia de la señalización por Fgf4 y Nodal en el mantenimiento *in vivo* de las células TS se hace evidente cuando se generan mutantes para el receptor *Fgfr2* o para las convertasas *Spc1* y *Spc4*. Los embriones mutantes para *Fgfr2* mueren en peri-implantación presentando defectos severos en el desarrollo del trofoblasto, mientras que los dobles mutantes para *Spc1* y *Spc4* muestran defectos asociados a la pérdida de señalización por Nodal en el epiblasto y son incapaces de mantener las células TS en el ExE (Arman et al., 1998; Guzman-Ayala et al., 2004).

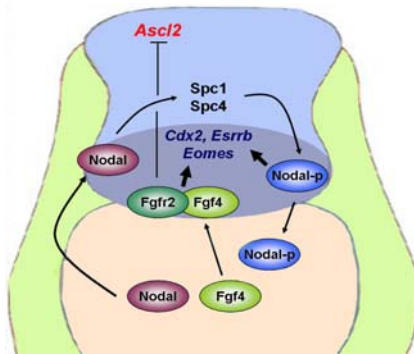


Figura 2. La señalización del epiblasto actúa en el mantenimiento de las células TS. Hay dos vías mediante las que el epiblasto contribuye al mantenimiento de las células TS. Nodal expresado en el epiblasto es procesado en el ExE por las convertasas Spc1 y Spc4. Nodal procesado (Nodal-p) contribuye a la expresión de *Cdx2*, *Esrrb* y *Eomes* en la población de células TS y al establecimiento del patrón embrionario en el epiblasto. Por otro lado, Fgf4 expresado en el epiblasto se une a su receptor Fgfr2 en el trofoblasto contribuyendo de igual modo al mantenimiento de la expresión de *Cdx2*, *Esrrb* y *Eomes* en el nicho de células TS, al tiempo que reprime la expresión de *Ascl2* que queda restringida al EPC.

Del mismo modo que la señalización por parte del epiblasto es necesaria en el mantenimiento de las células TS, señales originadas en el ExE son requeridas para el mantenimiento del patrón tanto en el epiblasto como en el endodermo visceral. Así por ejemplo, la señalización por parte del ExE ejerce un efecto inhibitorio sobre el endodermo visceral embrionario, de modo que restringe la expresión de *Cer1* y *Lefty1* al endodermo visceral distal y anterior (Rodriguez et al., 2005). En el estadio E5.5 Nodal expresado en el epiblasto induce la especificación del endodermo visceral distal, que migrará hacia un lado del embrión para constituir el endodermo visceral anterior (AVE, del inglés Anterior Visceral Endoderm). La expresión de *Cer1* y *Lefty1*, entre otros, resulta en la inhibición de la señalización por Nodal en el epiblasto, de modo que queda restringido al lado opuesto del embrión: la parte posterior donde comienza la especificación del mesodermo y formación de la línea primitiva, que dará lugar a la formación de las tres capas germinales que conforman el embrión (revisado en (Tam and Loebe, 2007)).

3 Expansión del trofoblasto y formación de la placenta

A los 7.5 días de desarrollo el ExE se ha expandido para formar el ectodermo del corion mientras que las células del EPC se han diferenciado para formar el espongiotrofoblasto. En la parte posterior del embrión se va a especificar ahora una población celular de carácter extraembrionario: el mesodermo del alantoides (Fig. 1). El alantoides emerge de la parte posterior embrionaria y establece contacto con el corion alrededor de los 8.5 días de desarrollo en un proceso que se conoce como la fusión corio-alantoidea. La fusión corio-alantoidea da lugar en estadios posteriores a la formación de la red de vasos sanguíneos encargados de la comunicación materno-fetal así como al cordón umbilical, generados ambos a partir del alantoides. La formación de los vasos sanguíneos feto-placentarios crea una densa red en esa región del trofoblasto, que conforma el denominado laberinto. En el laberinto encontramos los vasos sanguíneos rodeados fundamentalmente por un tipo de células multinucleadas llamadas sincitiotrofoblasto, acompañadas de algunas células gigantes y células mononucleadas. El espongiotrofoblasto queda como una capa de tejido estructural que rodea al laberinto y que se separa del tejido materno por células gigantes de trofoblasto (Fig. 1) (Rossant and Cross, 2001).

4 Células madre derivadas del blastocisto

Todas las estructuras tanto embrionarias como extraembrionarias generadas a los largo del desarrollo provienen de procesos de diferenciación que tienen su origen en las células que componen los linajes del blastocisto. Por éste motivo las células que forman el blastocisto se consideran células madre de cada uno de estos linajes, y se denominan así “células madre embrionarias” (células ES, el inglés Embryonic Stem cells) aquellas que forman el epiblasto, “células madre del trofoblasto” (células TS) aquellas de las que deriva el trofoblasto, y “células madre del endodermo extraembrionario” (células XEN, del inglés Extraembryonic Endoderm), de las que derivan el endodermo primitivo y parietal. Estos tipos celulares se pueden derivar del blastocisto y mantener en cultivo de manera indefinida, conservando el potencial para diferenciarse en cualquiera de los tipos celulares que conforman el tejido del que son precursoras (Roper and Hemberger, 2009). Constituyen así el modelo más representativo para el estudio *in vitro* de los procesos que tienen lugar en las etapas del desarrollo próximas al estadio de blastocisto.

4.1 Células madre de trofoblasto

Desde su derivación por primera vez, las células madre de trofoblasto o células TS han resultado de gran utilidad para el estudio *in vitro* de los procesos relacionados con el desarrollo de los linajes de la placenta. Se encuentran presentes en el embrión desde el estadio de blastocisto hasta aproximadamente los 7.5 días de desarrollo, y se han derivado satisfactoriamente tanto de blastocistos como de embriones en estadio E6.5, presentando en ambos casos propiedades muy similares. Son capaces de proliferar de manera indefinida y

contribuyen a la formación de todos los linajes del trofoblasto al ser incorporadas de nuevo al embrión en desarrollo mediante la formación de quimeras, cubriendo tanto la propiedad de auto-renovación como la de multipotencia necesarias para ser consideradas como células madre (Tanaka et al., 1998).

En cultivo las células TS tienen iguales propiedades que las presentes en el embrión *in vivo*. Así, para su mantenimiento requieren del aporte externo de Fgf4 y de Activina, factor análogo a Nodal, de forma que las vías de señalización activadas por estos dos factores contribuyen al mantenimiento de los factores de multipotencia entre los que se encuentran *Cdx2*, *Eomes* y *Esrrb* (Erlebacher et al., 2004; Rielland et al., 2008; Tanaka et al., 1998). La activación de Erk1/2 como consecuencia de la señalización por Fgf4 en las células TS es además necesaria para la inhibición de la proteína pro-apoptótica Bim, promoviendo de este modo la supervivencia celular (Yang et al., 2006).

En ausencia de Fgf4 o Activina las células TS se diferencian fundamentalmente en células gigantes de trofoblasto poliploides, pero también en células de sincitiotrofoblasto multinucleadas y de espongiotrofoblasto mononucleadas. Se han descrito algunos factores cuya desregulación promueve la diferenciación de las células TS hacia uno u otro tipo celular dentro del linaje del trofoblasto independientemente del aporte de Fgf4 o Activina. Así, la sobre-expresión de los factores de transcripción *Hand1*, *Stra13* o *Gcm1* promueve la salida del ciclo celular y diferenciación hacia células gigantes, en el caso de *Hand1* y *Stra13*, o hacia células de sincitiotrofoblasto, en el caso de *Gcm1*. Por su parte el factor de transcripción *Ascl2* parece promover cierta proliferación transitoria independiente de Fgf4 y Activina en células determinadas a diferenciarse en sincitiotrofoblasto y posteriormente en células gigantes (Hughes et al., 2004).

En su diferenciación hacia células gigantes de trofoblasto, las células TS pasan por distintas fases de endo-reduplicación del DNA mediante la cual se producen ciclos sucesivos de síntesis del DNA sin mitosis, adquiriendo así un carácter poliploide. Para que se de este fenómeno se ha de evitar el control que existe en el ciclo celular mediante el cual se impide la entrada de nuevo en fase de síntesis si no se ha producido la mitosis y que permite mantener las células del organismo diploides. Recientemente se ha descrito un mecanismo dependiente de Fgf4 mediante el cual las células TS pueden evitar este control del ciclo celular y endo-reduplicar su DNA diferenciándose a células gigantes. En este proceso están implicadas las proteínas inhibidoras de kinasa dependiente de ciclina Cdkn1c (p57/Kip2) y Cdkn1a (p21/Cip). En ausencia de Fgf4 se produce en las células TS un aumento de la expresión de estas dos proteínas. Mientras que p57 inhibe la función de la kinasa dependiente de ciclina Cdk1, necesaria para la entrada en mitosis, p21 suprime la actividad de Chk1, kinasa de control activada en respuesta al daño en el DNA. De este modo las células TS pueden evitar la entrada en mitosis sin ser eliminadas por los mecanismos de control del ciclo celular (Ullah et al., 2008).

REGULACIÓN POST-TRANSCRIPCIONAL MEDIADA POR miRNAs

Durante décadas el estudio de la regulación de la función génica durante el desarrollo se ha centrado fundamentalmente en la regulación a nivel transcripcional por parte de factores de transcripción. Sin embargo, si bien es cierto que los factores de transcripción juegan un papel fundamental en la regulación génica, no deben pasarse por alto todos los demás factores implicados en la generación última del producto de la expresión, ya sean proteínas o RNAs. En este sentido, durante la última década se ha abierto un nuevo foco de atención sobre el estudio de la regulación a nivel post-transcripcional por parte de RNAs no codificantes de pequeño tamaño, y más concretamente de los microRNAs (miRNAs).

En 1993 se describió en *Caenorhabditis elegans* por primera vez un RNA endógeno de pequeño tamaño, *lin-4*, capaz de reprimir de forma post-transcripcional la expresión un gen, *lin-14*, mediante su unión por secuencias complementarias al mRNA de éste. Esta interacción regulada entre *lin-4* y el mRNA de *lin-14* demostró ser además fundamental para la correcta especificación del destino celular durante el desarrollo larvario de *C.elegans* (Lee et al., 1993; Wightman et al., 1993). *lin-4* fue el primer miRNA descrito, sin embargo no fue hasta el año 2000, con el descubrimiento de *let-7* también en *C.elegans* y también implicado junto con *lin-4* en el desarrollo larvario (Reinhart et al., 2000), cuando comenzó a estudiarse de forma más exhaustiva la implicación de los miRNAs en la regulación génica.

Los miRNAs son RNAs de 20 a 22 nucleótidos de longitud que ejercen su acción mediante la unión complementaria a la región 3'UTR de RNAs mensajero diana, actuando como represores de la función génica a nivel post-transcripcional. Este mecanismo de interacción RNA-RNA es muy similar al que ejercen los RNAs pequeños de interferencia (siRNAs, del inglés, Short Interfering RNAs), RNAs generalmente exógenos que son procesados a RNAs de pequeño tamaño y actúan uniéndose de forma complementaria al mRNA y degradándolo. Se ha descrito la generación de siRNAs de forma endógena (endo-siRNAs), y parece que su función principal es en este caso el silenciamiento de secuencias repetitivas. Otros RNAs de pequeño tamaño que se unen al RNA provocando su degradación son los piRNAs, que se han descrito únicamente en la línea germinal, donde parecen tener un papel fundamental en el silenciamiento de transposones (Ghildiyal and Zamore, 2009).

De entre las distintas clases de RNAs de pequeño tamaño descritas, los miRNAs son de especial relevancia porque se expresan virtualmente en todos los tipos celulares y se encuentran formando parte activa de la regulación de muchos procesos incluyendo el desarrollo, proliferación celular, apoptosis, metabolismo y también en numerosas enfermedades entre las que se encuentra el cáncer (Kloosterman and Plasterk, 2006; Stefani and Slack, 2008; Ventura and Jacks, 2009).

1 Organización genómica de los miRNAs

Desde su descubrimiento en *C.elegans* se han descrito multitud de miRNAs tanto en animales como en plantas y virus, e incluso más recientemente en el alga unicelular *Chlamydomonas reinhardtii* (Molnar et al., 2007; Zhao et al., 2007). Muchas familias de miRNAs se encuentran altamente conservadas entre distintas especies animales (Lagos-Quintana et al., 2003), como es el caso de la familia *let-7*, conservada casi todos los organismos bilaterales (Pasquinelli et al., 2003; Pasquinelli et al., 2000). Sin embargo, esta elevada conservación no se extiende a las plantas, cuyos miRNAs no se encuentran conservados en el reino animal y presentan además un mecanismo de acción y una ruta de biogénesis que difieren significativamente de los que encontramos en animales (Chen and Rajewsky, 2007).

Los miRNAs se expresan a partir de precursores más largos (pri-miRNAs) codificados en el genoma. Las últimas estimaciones apuntan a un total de 533 loci codificantes para miRNAs descritos en el genoma humano, 442 en el genoma de ratón, 135 en *C.elegans*, 93 en *D.melanogaster* y 184 en *A.thaliana* (Griffiths-Jones et al., 2008). Los loci codificantes para miRNAs pueden encontrarse tanto en regiones intergénicas como en intrones, exones y UTRs de genes codificantes y no codificantes para proteínas (Griffiths-Jones et al., 2008; Lagos-Quintana et al., 2003; Rodriguez et al., 2004).

En el genoma podemos encontrar codificados frecuentemente miRNAs que distan menos de 10kb de otro miRNA vecino, de modo que forman lo que se denomina un cluster (Ambros, 2004; Bartel, 2004; Griffiths-Jones et al., 2008). El cluster de mayor tamaño descrito se encuentra en el dominio Dlk1-Gtl2, codificado en el cromosoma 14 de humanos y en su región homóloga del cromosoma 12 de ratón, y se extiende a lo largo de 40kb aproximadamente (Seitz et al., 2004). La proximidad dentro del cluster hace que en muchos casos varios miRNAs se transcriban de forma conjunta como un único precursor (pri-miRNA), que será más adelante procesado para dar los diferentes miRNAs maduros (Baskerville and Bartel, 2005), este es por ejemplo el caso del policistron miR-17-92 (He et al., 2005). Sin embargo, miRNAs pertenecientes a un mismo cluster pueden también estar regulados y transcritos de forma independiente como sucede en el locus miR-433-127 de ratón (Song and Wang, 2008).

2 Biogénesis de los miRNAs

Los miRNAs pasan por distintas fases de procesamiento desde su transcripción en el núcleo hasta su maduración e incorporación en el citoplasma en complejos ribonucleoproteicos, capaces de unirse a las regiones 3'UTR de mRNAs diana y provocar su degradación y/o impedir su traducción. La ruta de biosíntesis que siguen los miRNAs en vertebrados se encuentra esquematizada en la figura 3.

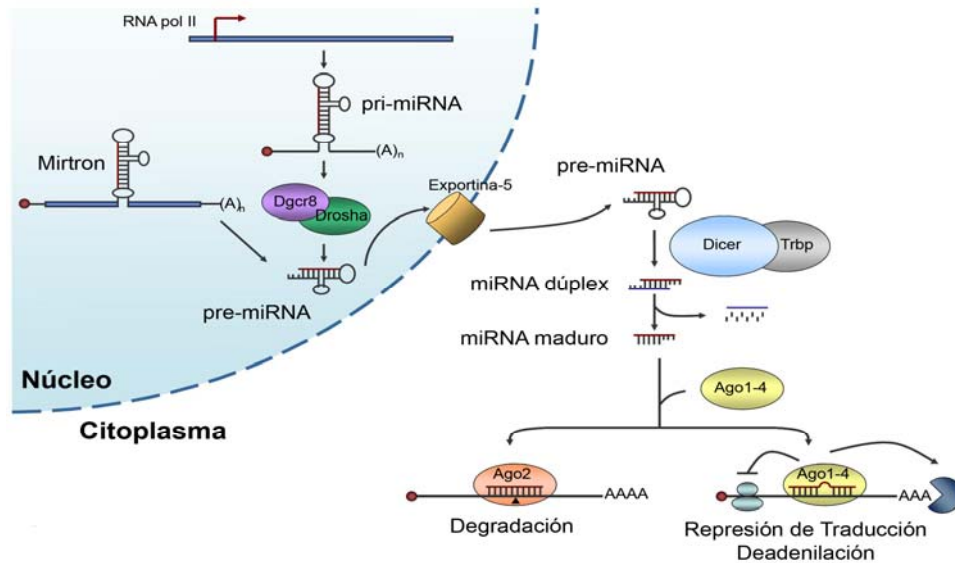


Figura 3. Ruta de biogénesis de los miRNAs en vertebrados. Los miRNAs se transcriben a pri-miRNAs por la RNA polimerasa II. Si éstos se encuentran localizados en un intrón se escinden directamente a partir del *splicing* del mRNA hospedador y pueden en algunos casos plegarse formando directamente un pre-miRNA o mirtron. Los pri-miRNAs son procesados en el núcleo por el complejo Drosha/Dgcr8 para dar lugar a precursores de ~ 70 nucleótidos de longitud, los pre-miRNAs. Éstos pre-miRNAs se exportan al citoplasma mediante la proteína de membrana Exportina-5, y una vez ahí son reconocidos por un complejo de proteínas entre las que se encuentran la RNasa de tipo III Dicer y la proteína Trbp, que estabiliza a Dicer. Dicer procesa el pre-miRNA generando un RNA de doble cadena de 20-22 nucleótidos de longitud. Este dúplex de RNA se separa del complejo de Dicer y una de las dos hebras es degradada. La hebra que permanece intacta constituye el miRNA maduro y es integrada en el complejo RISC (del inglés, *RNA-Induced Silencing Complex*), constituido por diversas proteínas entre las cuales juegan un papel fundamental cuatro miembros de la familia Argonauto (Ago1-4). El miRNA guía al complejo RISC hasta el mRNA diana, al que se une por complementariedad de bases. Si la unión miRNA-mRNA es perfectamente complementaria se produce la degradación del mRNA mediada por Ago2. Si por el contrario la unión es imperfecta se produce una represión de la traducción, acompañada en muchos casos de la deadenilación del mRNA, que acaba degradándose (modificado de (Filipowicz et al., 2008)).

2.1 Transcripción

Estudios llevados a cabo sobre la transcripción de los miRNAs apuntan a una regulación del proceso muy similar a la de los genes codificantes para proteínas convencionales. Aunque se han descrito casos de transcripción de pri-miRNAs por la RNA polimerasa III (Borchert et al., 2006), la mayoría se transcriben utilizando la RNA polimerasa II, que genera un pri-miRNA de longitud variable, que sufre un proceso de poliadenilación en el extremo 3' y adición de la "caperuza" o "cap" en el extremo 5' de forma similar a como ocurre en RNAs codificantes para proteínas (Lee et al., 2004). La transcripción de los pri-miRNAs está también sujeta a procesos de regulación similares a los de genes codificantes convencionales, tanto a nivel de regulación epigenética por modificación de la cromatina como a nivel de control por parte de factores de transcripción. Así, se ha visto que procesos como metilación del DNA y modificación de histonas pueden afectar a la activación o represión de la transcripción de los pri-miRNAs (Bueno et al., 2008; Lujambio and Esteller, 2009). Del mismo modo, se han descrito factores de transcripción que controlan la expresión de los miRNAs. Ejemplos de este tipo de regulación son *c-Myc* o *p53*, que controlan positiva o negativamente la expresión de distintos miRNAs (He

et al., 2007; O'Donnell et al., 2005). Por último, aquellos miRNAs que se encuentran integrados en los intrones de genes codificantes se pueden transcribir conjuntamente con dicho gen, de forma que comparten todos sus elementos reguladores (Rodríguez et al., 2004).

2.2 Procesamiento del pri-miRNA

Una vez transcritos, los pri-miRNAs se pliegan formando estructuras en horquilla de RNA bicatenario y son así procesados en el núcleo por el complejo proteico conocido como microprocesador, en el que intervienen la proteína Dgcr8, que une el pri-miRNA, y la enzima RNasa de tipo III Drosha, que lo corta generando horquillas de RNA de doble cadena de aproximadamente 70 nucleótidos de longitud denominadas pre-miRNAs (Denli et al., 2004).

La función del microprocesador se encuentra a su vez regulada de diversas formas (Winter et al., 2009) y existen además miRNAs cuyo procesamiento no requiere de la acción del microprocesador. Es el caso de algunos de los miRNAs localizados en intrones de genes codificantes. Estos se transcriben junto con su gen hospedador y se liberan del transcrito del gen durante el proceso de *splicing* de éste. Si el intrón liberado es del tamaño adecuado puede plegarse formando directamente un pre-miRNA que no requiere de la acción del microprocesador. Los miRNAs así generados se denominan mirtrones (Berezikov et al., 2007; Ruby et al., 2007).

2.3 Procesamiento del pre-miRNA

Los pre-miRNAs generados por el microprocesador o provenientes de mirtrones son reconocidos en el núcleo por la proteína de membrana Exportina-5, que los transfiere al citoplasma (Yi et al., 2003). Una vez en el citoplasma la RNasa de tipo III Dicer se encarga del procesamiento de todos los pre-miRNAs para generar un RNA de doble cadena de 20-22 nucleótidos de longitud (Bernstein et al., 2001; Ketting et al., 2001). Dicer actúa en conjunto con la proteína Trbp, que aunque no es necesaria para el procesamiento lo facilita mediante la estabilización de Dicer (Chendrimada et al., 2005; Melo et al., 2009).

El dúplex de RNA generado por Dicer se disocia de este complejo proteico y una de las hebras es degradada, mientras que la otra, que constituye el miRNA maduro, es incorporada en el complejo proteico RISC (del inglés RNA-Induced Silencing Complex). La degradación de una hebra u otra depende en cierta medida de la estabilidad del apareamiento en el extremo 5' de cada una (Khvorova et al., 2003). En muchos casos se han detectado por técnicas de secuenciación de RNA de pequeño tamaño miRNAs procedentes de ambas hebras, lo que indica que ambos podrían ser funcionales. En estos casos los miRNAs se nombran añadiendo el sufijo -3p o -5p en cada caso (por ejemplo miR-140-3p y miR-140-5p) (Griffiths-Jones et al., 2008).

Los componentes principales del complejo RISC, aunque no los únicos, son las proteínas de la familia Argonauta, que son las efectoras de la función represora de los miRNAs. En mamíferos hay cuatro proteínas Argonauta (Ago1-4) que funcionan en la maquinaria de represión mediada

por RNA de interferencia, de las cuales únicamente Ago2 tiene capacidad para degradar directamente el RNA (Liu et al., 2004; Lu et al., 2005). El miRNA maduro incorporado en el RISC se une específicamente por complementariedad de bases al mRNA diana, sirviendo así como guía a las proteínas del complejo RISC, que mediarán la degradación del mRNA y/o impedirán su traducción (Filipowicz et al., 2008).

2.4 Represión mediada por miRNAs

La represión mediada por miRNAs se produce a través de su unión a las regiones 3'UTR de los mRNA diana. Aunque se han descrito miRNAs cuyas dianas se encuentran en regiones codificantes o en los extremos 5'UTR (Tay et al., 2008), parece que en la mayoría de los casos no resultan dianas muy eficientes (Baek et al., 2008; Grimson et al., 2007).

El mecanismo por el cual los miRNAs ejercen su acción sobre el mRNA diana depende en gran medida de la especificidad de la unión entre ambos. La región más importante del miRNA y que va a determinar la especificidad de la unión son 7-8 nucleótidos situados en posición 5', lo que se denomina la región "semilla", mientras que la especificidad de la unión de los nucleótidos situados en 3' coopera aportando una mayor estabilidad termodinámica que contribuye a la eficacia de la acción del miRNA (Grimson et al., 2007).

La perfecta complementariedad en el apareamiento del miRNA tiene como consecuencia la degradación del mRNA mediada por la acción de Ago2. Este es el mecanismo de degradación que tiene lugar en el caso de los siRNAs y de los miRNAs en plantas. Sin embargo, en animales la unión del miRNA a su diana es sólo completamente complementaria en su región "semilla", y se estabiliza la unión en función del grado de complementariedad de la región 3' (Grimson et al., 2007). Esto hace que el efecto del miRNA sobre su mRNA diana sea de impedimento de la traducción de éste a proteína en lugar de su degradación. A pesar de que la unión imperfecta del miRNA a su diana tiene como principal consecuencia el impedimento de la traducción de éste actuando a distintos niveles (Filipowicz et al., 2008), recientemente se ha visto que este tipo de interacción entre el miRNA y el mRNA provoca en la mayor parte de los casos el reclutamiento de proteínas que degradan progresivamente la cola de poli(A) del mRNA seguido del resto de la molécula, lo que en definitiva tiene como consecuencia la degradación del mRNA por un mecanismo diferente al implicado en las uniones perfectamente complementarias. Sin embargo, para un mRNA diana concreto la represión dada por uniones imperfectas va a ser en la mayor parte de los casos mayor a nivel de proteína que a nivel de mRNA, y en ciertos casos ésta última es prácticamente inapreciable (Baek et al., 2008; Filipowicz et al., 2008; Selbach et al., 2008; Zipprich et al., 2009).

A pesar de que la gran mayoría de los estudios sobre miRNAs describen su actuación en el sentido indicado más arriba, se han descrito casos de miRNAs que provocan un aumento en la traducción del mRNA (Vasudevan et al., 2007), miRNAs que son importados del citoplasma al núcleo (Hwang et al., 2007), o mRNAs que acortan selectivamente sus 3'UTR para eliminar sitios de unión de miRNAs determinados (Legendre et al., 2006; Sandberg et al., 2008).

FUNCIÓN DE LOS miRNAs EN EL DESARROLLO

La descripción de los miRNAs como mecanismo de regulación de la expresión génica es relativamente reciente, por lo que las aproximaciones al estudio de la función de los miRNAs pasan en primer lugar por su detección en distintos órganos, tejidos y tipos celulares. A continuación, es necesario saber qué mRNAs pueden estar sujetos a regulación por parte de miRNAs. Éstos pueden unirse potencialmente a multitud de mRNAs diana mediante la complementariedad en la región “semilla”, sin embargo no todas las secuencias complementarias a dicha región constituyen sitios de unión de funcionales. Por ello se han de utilizar diferentes métodos de predicción para detectar potenciales sitios reales de unión de miRNAs. Por último, la funcionalidad de los sitios predichos para un determinado miRNA ha de comprobarse experimentalmente, mediante la realización de distintos ensayos tanto *in vitro* como *in vivo*.

A continuación se describen las diferentes aproximaciones que se están llevando a cabo para el estudio de la función de los miRNAs, y que hasta el momento han permitido situarlos como importantes reguladores de diferentes procesos. Asimismo, centraremos nuestra atención sobre la implicación que los miRNAs tienen sobre el desarrollo embrionario.

1 Análisis de expresión de los miRNAs y búsqueda de mRNAs diana

1.1 Análisis de expresión

El descubrimiento de nuevos miRNAs se basó en un principio en la combinación de métodos experimentales, como es la generación de librerías de RNAs de pequeño tamaño y su posterior secuenciación (Lagos-Quintana et al., 2003; Lagos-Quintana et al., 2002; Landgraf et al., 2007), y métodos de predicción computacionales, basados fundamentalmente en la búsqueda de secuencias homólogas a miRNAs conocidos y con capacidad potencial para formar una horquilla de doble cadena de RNA de propiedades similares a las que forman los miRNAs descritos (Bartel, 2004; Lim et al., 2003; Seitz et al., 2004). Más recientemente, el desarrollo de las técnicas de secuenciación de alto rendimiento (High –Throughput Sequencing) está permitiendo la detección de un número mayor de miRNAs a partir de librerías de RNAs de pequeño tamaño. Todos los miRNAs detectados se han incorporado a la base de datos miRBase (<http://microrna.sanger.ac.uk/>).

La construcción de librerías de RNAs de pequeño tamaño y su posterior secuenciación ha resultado ser de gran utilidad a la hora de detectar miRNAs específicamente expresados en determinados tejidos, órganos o tipos celulares (Landgraf et al., 2007). Asimismo, se han desarrollado técnicas de microarray de miRNAs con el fin de obtener perfiles de expresión específicos de tejido (Baskerville and Bartel, 2005). Esta técnica permite analizar la expresión

de todos los miRNAs descritos y predichos en un determinado tejido, sin embargo requiere de una posterior validación mediante Northern o PCR cuantitativa.

La descripción del patrón de expresión de los miRNAs estuvo durante un tiempo limitada por la dificultad en la realización de hibridaciones *in situ* dado el pequeño tamaño de los miRNAs, lo que empujó a ciertos investigadores al diseño de métodos alternativos. Así, tanto en *Drosophila* como en ratón se desarrolló una técnica basada en la generación de embriones transgénicos para un gen reportero constitutivo que posee en su región 3'UTR dianas para un miRNA determinado. Esto permite la detección de la expresión del gen constitutivo en todos los tejidos salvo en aquellos en los que se expresa el miRNA que los reprime (Brennecke et al., 2003; Mansfield et al., 2004). Más recientemente se han perfeccionado las técnicas de hibridación *in situ* para la detección de miRNAs, mediante el diseño de sondas artificiales que portan una modificación que las hace altamente estables y específicas. Estas son las denominadas sondas LNA (del inglés, Locked Nucleic Acid) (Kloosterman et al., 2006).

1.2 Búsqueda de mRNAs diana

Una vez conocido el patrón de expresión de miRNAs determinados, es necesario saber qué genes pueden estar regulando. Esto no resulta fácil si tenemos en cuenta que un miRNA va a ejercer su función sobre los mRNAs diana dependiendo fundamentalmente del grado de complementariedad que presenten en los 7-8 nucleótidos de la región “semilla”. Esto hace que sea muy sencillo para un mRNA generar un sitio de unión para un determinado miRNA en su región 3'UTR, lo que se refleja en el elevado número de mRNAs que es capaz de regular un único miRNA, del orden de cientos en la mayoría de los casos (Lim et al., 2005).

Durante los últimos años se ha invertido mucho esfuerzo en la generación de modelos de predicción de mRNAs diana para cada miRNA, basados no sólo en la complementariedad de la región “semilla” sino también otros parámetros como la estabilidad termodinámica de la unión, a la que contribuye la región en 3' del miRNA; el contenido en AU; el grado de conservación de los sitios de unión de los miRNAs; el número de sitios y la localización de los mismos en cada 3'UTR (Grimson et al., 2007). Se han desarrollado diferentes programas para la predicción de mRNAs diana basados en algoritmos que tienen en cuenta todos estos parámetros. Entre ellos encuentran miRanda (www.microrna.org, (Betel et al., 2008)), PicTar (www.pictar.mdc-berlin.de, (Krek et al., 2005)) y TargetScan (www.targetscan.org, (Grimson et al., 2007)). El fundamento del algoritmo en que se basan las predicciones de cada uno de estos programas se encuentra revisado en (Maziere and Enright, 2007).

2 Implicación de los miRNAs en el desarrollo embrionario

El descubrimiento de los primeros miRNAs en *C. elegans* (*lin-4* y *let-7*) se llevó a cabo por procedimientos genéticos clásicos: análisis de genes cuyas mutaciones presentan un fenotipo. En este caso el fenotipo observado implicaba una función de estos miRNAs en el control

temporal de la especificación de linajes durante el desarrollo larvario del gusano (Lee et al., 1993; Reinhart et al., 2000; Wightman et al., 1993). Desde entonces se ha demostrado la importancia de los miRNAs en multitud de procesos relacionados con el desarrollo (Stefani and Slack, 2008). Prueba de ello es el estudio llevado a cabo por Stark y colaboradores (2005), en el que clasifican los genes en dos tipos: “anti-dianas” y “dianas”. El primer caso se refiere a genes situados bajo presión selectiva de forma que evitan ser dianas de regulación por parte de miRNAs (por ejemplo, manteniendo las regiones 3'UTR más cortas y con menor densidad de sitios de unión de miRNAs); este tipo de genes se expresan en general de forma ubicua y están envueltos en procesos celulares básicos. Por otro lado, los genes “diana” presentan en general regiones 3'UTR más largas y mayor densidad de sitios de unión de miRNAs, tienen en muchos casos patrones de expresión específicos de tejido y se encuentran sobre-representados en procesos relacionados con el desarrollo (Stark et al., 2005).

Actualmente se manejan tres modelos básicos de actuación de los miRNAs durante el desarrollo: encendido/apagado de genes, degradación de mensajeros que ya no son necesarios y mantenimiento de patrones de expresión previamente establecidos (Fig. 4) (Plasterk, 2006; Shkumatava et al., 2009; Stark et al., 2005).

En el mecanismo de encendido/apagado de genes los miRNAs actuarían de forma similar a como lo hacen los factores de transcripción, determinando por sí mismos dónde se va a expresar un gen concreto y formando parte integrante del evento del desarrollo. En este caso el miRNA y su gen diana tienen la expresión dirigida al mismo territorio, pero el miRNA se encarga de silenciar completamente la expresión del gen. En este tipo de regulación es en la que la mutación en un miRNA causa un fenotipo debido a la sobre-expresión del mRNA diana. Ejemplo de ello es la función de miR-61 en el desarrollo de *C. elegans*. En este caso el destino de las células de la vulva está determinado por la expresión de dos genes, de forma que la expresión del gen que determina el destino celular primario promueve la expresión de miR-61, que a su vez reprime al gen determinante del destino celular secundario (Fig. 4A) (Yoo and Greenwald, 2005).

En el segundo tipo de regulación, los miRNAs actuarían degradando aquellos transcritos que ya no son necesarios en situaciones como pueden ser la diferenciación en un determinado tipo celular o la transición materno-zigótica durante las primeras etapas del desarrollo embrionario. Un ejemplo de este tipo de regulación se ha descrito en el pez cebra, donde la expresión de la familia de miRNAs miR-430 comienza con la activación del genoma zigótico y actúa colaborando en la degradación de cientos de mRNAs mensajeros de origen materno, favoreciendo así la transición de expresión materno-zigótica (Fig. 4B) (Giraldez et al., 2006).

Por último está el tipo de regulación por el cual mecanismos como factores de transcripción ejercen la principal represión sobre los genes diana, y los miRNAs controlan los niveles de fondo de expresión que pudieran escapar al control previo contribuyendo así al mantenimiento de un patrón previamente establecido. Este tipo de regulación se puede dar en dos situaciones

(Fig. 4C). En un primer escenario los niveles de expresión finales del gen diana son elevados en el territorio en el que se expresa el miRNA que lo regula, sin embargo aquí la función del miRNA no está dirigida a reprimir de forma drástica al mRNA diana sino al control óptimo de la cantidad final de proteína. En una segunda situación los mensajeros diana se encuentran expresados a niveles muy bajos en los territorios de expresión del miRNA. Este es el caso por ejemplo, miR-1 y miR-124, que se encuentran altamente expresados en poblaciones concretas del músculo y el cerebro respectivamente, y sus respectivos mensajeros diana presentan niveles de expresión muy bajos en esos territorios (Lim et al., 2005). Otro ejemplo de este tipo se da en situaciones en las que el miRNA se expresa en territorios adyacentes al mRNA diana y en este caso contribuye a delimitar de forma más fina el patrón de expresión del gen diana. Este es el caso de miR-138 en la regulación de *cspg2* directa e indirectamente en la región adyacente al territorio de expresión de este gen durante el desarrollo del corazón en pez cebra (Morton et al., 2008).

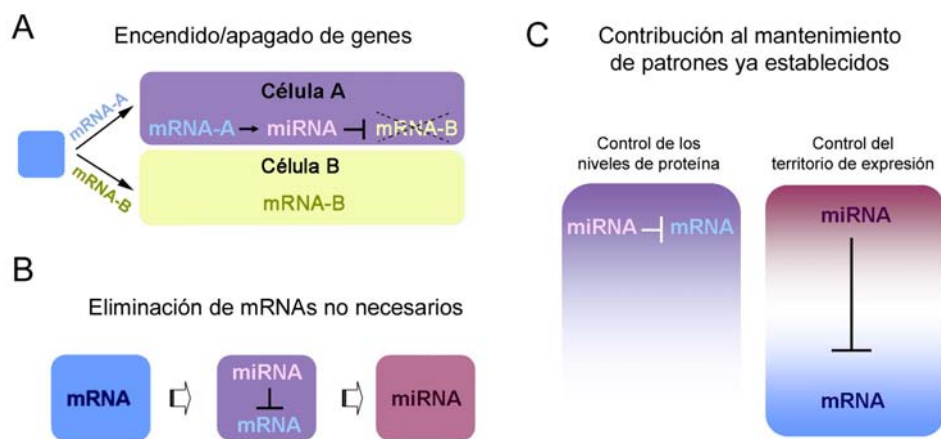


Figura 4. Mecanismos de actuación de los miRNAs durante el desarrollo. Los miRNAs pueden actuar de forma similar a como lo hacen los factores de transcripción, eliminando por completo un mRNA expresado en un determinado tipo celular y determinando así el destino de esa célula (A). Otra forma de actuación puede ir encaminada a la eliminación de mRNAs que ya no son necesarios. Así por ejemplo, durante un proceso de diferenciación se enciende la expresión de miRNAs que contribuyen a la eliminación del mRNA presente en la célula antes de la diferenciación y que ya no se requieren (B). Los miRNAs durante el desarrollo pueden contribuir a reforzar patrones de expresión previamente establecidos por otros medios como son los factores de transcripción. En este caso puede darse una expresión completamente solapante del mRNA y el miRNA que lo regula sin que esto suponga el apagado de la expresión génica. Aquí el miRNA estaría controlando de forma precisa los niveles finales de proteína de modo que sean óptimos. En otra situación el patrón de expresión de miRNA puede ser mutuamente exclusivo o sólo parcialmente solapante con el de su mRNA diana, y de este modo el miRNA contribuye a delimitar el territorio de expresión del mensajero. El territorio de expresión del mRNA se representa en color azul, el de expresión del miRNA en rojo, y el territorio en el que mRNA y miRNA están co-expresados se representa en color morado.

Este último tipo de regulación es el que se da más frecuentemente durante el desarrollo e implica una función de los miRNAs en conferir un grado extra de solidez a las redes regulatorias que realmente están estableciendo el destino celular (Hornstein and Shomron, 2006; Shkumatava et al., 2009; Stark et al., 2005). Dada esta forma de actuación, en muchos

casos la delección de miRNAs individuales no va a tener consecuencias fenotípicas fácilmente apreciables, por lo que el análisis de su función de forma individual resulta complejo. Esto se puso de manifiesto con la mutación sistemática de prácticamente todos los miRNAs presentes en *C.elegans*, que demostró que en un 83% de los casos la delección de un único miRNA no tiene un efecto fenotípico acusado (Miska et al., 2007). Existen además otros factores que dificultan el estudio funcional de miRNAs concretos. En primer lugar se encuentra el hecho de que un único miRNA puede regular cientos de mensajeros diana, de modo que su sobre-expresión o inhibición va modular la expresión de multitud de genes (Lim et al., 2005). Sin embargo, esto no implica necesariamente que la variación en todos los genes diana sea la causa de un determinado fenotipo asociado a un miRNA concreto. Así por ejemplo, a pesar de que *lin-4* tiene multitud de mRNAs diana, únicamente *lin-14* y en menor medida *lin-28* son responsables del fenotipo asociado a la pérdida de este miRNA en *C. elegans* (Ambros, 2004). Por otro lado, un mensajero puede presentar más de un sitio de unión para el mismo miRNA o sitios de unión para varios miRNAs diferentes que tendrán un efecto de represión multiplicativo (Tay et al., 2008). La posibilidad de regulación de un mensajero por varios miRNAs tiene un efecto de redundancia en la función de éstos, que se ve acrecentada por el hecho de que los miRNAs se encuentran en muchos casos formando parte de familias. Éstas se forman en su mayoría por duplicación génica, pudiendo los miembros que las forman mantenerse unidos formando un cluster o dispersarse por el genoma; y se caracterizan por tener la misma región “semilla”, lo que hace que diferentes miembros de una misma familia regulen prácticamente los mismos mRNAs (Chen and Rajewsky, 2007; Grimson et al., 2007).

Dadas las dificultades asociadas en muchos casos al estudio de la función de miRNAs concretos, se ha recurrido de forma común a la delección de proteínas implicadas en su ruta de biosíntesis, de forma que se impide la producción de todos ellos. La delección de todos los miRNAs de esta manera provoca un fenotipo notable, que se puede utilizar como base para el estudio posterior de miRNAs individuales que puedan estar relacionados con él. Para el estudio de procesos implicados en el desarrollo en ausencia de miRNAs se han generado líneas mutantes para las proteínas *Dgcr8*, *Dicer* y *Ago2* (Bernstein et al., 2003; Giraldez et al., 2005; Ketting et al., 2001; Morita et al., 2007; Wang et al., 2007), sin embargo la mayor parte de estudios se han centrado en el análisis tanto *in vivo* como *in vitro* de la falta de *Dicer*.

3 Mutantes de *Dicer*: modelo para el estudio de la función de los miRNAs en el desarrollo del ratón

En ratón los mutantes completos de *Dicer*, *Dgcr8* y *Ago2* presentan un fenotipo muy similar de letalidad embrionaria temprana (Bernstein et al., 2003; Morita et al., 2007; Wang et al., 2007). Asimismo, la generación de mutantes condicionales para *Dicer* o *Dgcr8* en epidermis así como en células ES demostró que el fenotipo obtenido en ambos casos es el mismo (en el caso de la epidermis) o muy similar (en el caso de las células ES) (Kanellopoulou et al., 2005; Wang et al.,

2007; Yi et al., 2009). Por este motivo se considera que los efectos observados en estos mutantes se pueden atribuir mayoritariamente a la falta de miRNAs. Sin embargo, la delección de *Dicer* conlleva la ausencia completa de miRNAs maduros, cosa que no sucede en el caso de los mutantes para *Dgcr8*, que aún conservarían aquellos miRNAs independientes del procesamiento por el complejo Drosha/Dgcr8, como son los mirtrones (Babiarz et al., 2008). Es por ello que el mutante de *Dicer* ha resultado el modelo de elección en la mayor parte de los casos para el estudio del desarrollo en ausencia de miRNAs.

Existen varias líneas de ratón mutantes para *Dicer*, tanto de pérdida completa de función como hipomórficas (Bernstein et al., 2003; Cobb et al., 2005; Harfe et al., 2005; Mudhasani et al., 2008; Yang et al., 2005). Los alelos que provocan la pérdida completa de función son aquellos en los que se interfiere con cualquiera de los dos dominios RNasa III de la proteína, ya que ambos son necesarios para el procesamiento de los miRNAs, o con el dominio PAZ de unión a pre-miRNAs. Bernstein y colaboradores describieron en 2003 por primera vez el fenotipo de falta total de función de *Dicer* en el ratón. El fenotipo descrito en este trabajo muestra una parada del desarrollo y muerte alrededor de los 7.5 días de desarrollo, acompañada de la bajada drástica de expresión del marcador de pluripotencia embrionaria *Pou5f1* (*Oct4*) en el epiblasto y la ausencia del marcador de mesodermo *T* (*Brachyury*) (Bernstein et al., 2003). Las demás líneas mutantes completas para *Dicer* generadas presentan el mismo fenotipo temprano, coincidiendo en gran medida con el observado en los mutantes para las proteínas *Dgcr8* y *Ago2* (Harfe et al., 2005; Morita et al., 2007; Mudhasani et al., 2008; Wang et al., 2007). Ninguno de estos mutantes ha sido analizado en profundidad, por lo que se desconoce en qué medida afecta la falta de miRNAs al desarrollo de los distintos linajes que componen el embrión temprano de ratón. Sin embargo, el hecho de que sea posible derivar células ES mutantes para *Dicer* o *Dgcr8* indica que el fenotipo observado en los mutantes no se debe a la ausencia de células madre embrionarias (Kanellopoulou et al., 2005; Murchison et al., 2005; Wang et al., 2007).

Análisis llevados a cabo en ratón con mutantes de *Dicer* condicionales en los que la delección se lleva a cabo en tejidos específicos durante el desarrollo parecen indicar que no existen defectos de establecimiento de patrón asociados a la falta de miRNAs. Así, la delección específica de *Dicer* en el mesodermo de la extremidad en estadio E9.5 no afecta a la diferenciación de los distintos tipos celulares, pero provoca no obstante una disminución notable en el tamaño de la extremidad asociada a un incremento de apoptosis sin que la proliferación se vea comprometida (Harfe et al., 2005). Fenotipos similares de muerte celular masiva y defectos en morfogénesis se observan en los mutantes condicionales de *Dicer* tanto en músculo esquelético como en pulmón (Harris et al., 2006; O'Rourke et al., 2007).

3.1 Función de los miRNAs en el desarrollo embrionario temprano *in vivo*

Mediante el uso de mutantes de *Dicer* condicionales y en algunos casos mediante ensayos de función de miRNAs específicos, se ha abordado el estudio de la función de los miRNAs durante

la organogénesis. Sin embargo, se conoce poco acerca de los mecanismos de regulación que estos represores ejercen en etapas tempranas del desarrollo.

Dos estudios demostraron la presencia de miRNAs de origen materno así como su importancia en la ovogénesis y la primera división del embrión de ratón. Así, la delección de *Dicer* en oocitos maduros tiene como consecuencia la formación de husos mitóticos desorganizados, afectando al progreso a través de la primera división del embrión (Murchison et al., 2007; Tang et al., 2007). El análisis de los perfiles de expresión de miRNAs en el embrión temprano de ratón se reduce a estadios de hasta 8 células. Los resultados de este análisis mostraron que la familia más abundante en el oocito y cigoto es *let-7*, mientras que el cluster más abundante es miR-17-92. Los miRNAs de origen materno son degradados en su mayoría con la activación del genoma cigótico, que tiene lugar en estadio de 2 células, y ahí comienza la transcripción activa por parte de éste de miRNAs. En este caso el cluster cuya expresión se activa más rápidamente en el embrión es el miR-290 (Tang et al., 2007). Los patrones de expresión de miRNAs en etapas tempranas de post-implantación no han sido descritos, con la única excepción de varios miembros de la familia de miR-17 y el miRNA miR-302, homólogo a la familia miR-430 de pez cebra y miR-427 de xenopus (Card et al., 2008; Foshay and Gallicano, 2009). La expresión de los miembros de la familia miR-17 se ha relacionado con tejidos diferenciados en el embrión temprano como son el trofoectodermo y endodermo primitivo del blastocisto tardío (éste es el caso de miR-17-5p y miR-20), y con las poblaciones de mesodermo nacientes durante la gastrulación, que expresan miR-17-5p y miR-93 (Foshay and Gallicano, 2009). En cuanto a miR-302, se ha detectado su expresión específica en el epiblasto de embriones en estadio E6.5 (Card et al., 2008).

Estas breves descripciones de patrones de expresión son los únicos datos que hasta el momento se tienen acerca de la expresión de miRNAs en el embrión temprano de ratón. Del mismo modo, no se ha abordado el estudio del desarrollo embrionario durante estas etapas en ausencia de miRNAs. En su lugar, la mayor parte de los trabajos se han llevado a cabo en modelos *in vitro* utilizando células madre embrionarias.

3.2 Función de los miRNAs *in vitro*: análisis en células madre embrionarias

En lo que respecta al estudio de la función de los miRNAs, únicamente las células madre correspondientes al linaje embrionario (ES) han sido analizadas. El fenotipo de falta de función de miRNAs en células ES se ha estudiado mediante el análisis de células mutantes para *Dicer* o *Dgcr8* (Calabrese et al., 2007; Kanellopoulou et al., 2005; Murchison et al., 2005; Wang et al., 2007). La delección de cualquiera de estas dos proteínas parece tener un efecto específico sobre el procesamiento de los miRNAs sin que otros RNAs de pequeño tamaño se vean afectados (Babiarz et al., 2008). Asimismo y de modo similar a como sucede en los embriones, ambos tipos de mutantes presentan un fenotipo muy parecido en células ES, resultante en una drástica bajada de proliferación e incapacidad para diferenciarse, por lo que éste se atribuye fundamentalmente a la ausencia de miRNAs maduros (Kanellopoulou et al., 2005; Wang et al.,

2007). Los perfiles de expresión de miRNAs en células ES indican que los más abundantes son aquellos que se han relacionado con la regulación del ciclo celular y con procesos de oncogénesis (Calabrese et al., 2007), por lo que parece que gran parte de la maquinaria de interferencia por miRNAs en este tipo celular está dirigida al control de este tipo de procesos. Entre estos miRNAs se encuentran miR-21 y los clusters miR-17-92, miR-15b-16 y miR-290-295, éste último descrito como específico de células ES y embrión en pre-implantación (Houbaviy et al., 2005; Houbaviy et al., 2003).

El defecto de diferenciación asociado a la falta de miRNAs se ha estudiado principalmente usando las células ES mutantes para *Dicer* (Sinkkonen et al., 2008). La ausencia del cluster miR-290 se ha descrito como la principal causa del defecto de diferenciación de las células ES, que en ausencia de estos miRNAs son incapaces de reprimir el gen de pluripotencia *Pou5f1* (Oct4). Los miRNAs de este cluster reprimen al mensajero del gen *Rbl2* (retinoblastoma like 2), que a su vez inhibe a las DNA metiltransferasas Dnmt1, Dnmt3a y Dnmt3b, encargadas del silenciamiento por metilación de, entre otros, *Pou5f1* (Oct4) (Benetti et al., 2008; Sinkkonen et al., 2008). Los defectos de proliferación de las células ES en ausencia de miRNAs se han estudiado en las células mutantes para *Dgcr8*. Se ha implicado en este caso a varios clusters de miRNAs que regulan la transición de fase G1 a S del ciclo celular actuando a diferentes niveles. Entre ellos se encuentra de nuevo el cluster miR-290, que reprime a varios miembros implicados en la salida del ciclo celular como son *Rbl2*, *Lats2* o *Cdkn1a* (p21) (Wang et al., 2008).

La regulación por miRNAs en células ES parece integrarse dentro de la red génica que controla la pluripotencia embrionaria, y en la que se encuentran factores de transcripción fundamentales para el mantenimiento de ésta como son *Nanog*, *Oct4* y *Sox2* (Marson et al., 2008). Así, se ha visto por un lado que miRNAs cuya expresión aumenta con la diferenciación de las células ES se encuentran regulando directamente estos genes de pluripotencia (Tay et al., 2008); y por otro lado *Nanog*, *Oct4* y *Sox2* se asocian a promotores de miRNAs altamente expresados en células madre embrionarias (Card et al., 2008; Marson et al., 2008).

En este trabajo hemos abordado el estudio de la función de los miRNAs en el desarrollo embrionario temprano centrándonos en el linaje del trofoblasto. La caracterización de embriones y células mutantes para *Dicer*, en combinación con el análisis de los perfiles de expresión de miRNAs en embriones en estadios de post-implantación y en células TS, nos ha permitido describir el requerimiento de los miRNAs para el mantenimiento de las células madre del trofoblasto tanto *in vivo* como *in vitro*, así como la búsqueda de miRNAs candidatos a contribuir a la regulación génica durante el desarrollo temprano este linaje.

Objetivos

1. Estudio de la función de los miRNAs durante el desarrollo temprano del trofoblasto del embrión de ratón.

Para ello se estudiará el fenotipo en el trofoblasto de embriones mutantes completos para la proteína Dicer, encargada de la maduración de los miRNAs.

2. Descripción del fenotipo causado por la ausencia de miRNAs en células madre de trofoblasto (células TS) en cultivo.

Este objetivo incluye la derivación de una línea de células TS condicionales para *Dicer* y el análisis de su fenotipo tras la delección del gen *in vitro*.

3. Caracterización de la expresión de miRNAs en células TS y en embriones de ratón en estadios tempranos de post-implantación.

Este análisis se llevará a cabo mediante la realización de un array de expresión de miRNAs que se combinará con la descripción mediante hibridación *in situ* de miRNAs en el embrión temprano.

4. Análisis de los mecanismos de acción de miRNAs en células TS.

Para ello se combinarán los resultados obtenidos del análisis del fenotipo de falta de función de Dicer en células TS (Objetivo 2) con los datos de expresión de miRNAs obtenidos del análisis de los perfiles de expresión (Objetivo 3).

Materiales y Métodos

ANÁLISIS *IN VIVO* DEL FENOTIPO DE FALTA DE FUNCIÓN DE DICER EN EL TROFOBlasto

1 Líneas de ratón y genotipaje

La línea de ratón *Dicer*^{+/-} se obtuvo del cruce de ratones *Dicer*^{fx/fx} (Cobb et al., 2005) con ratones *Sox2Cre*^{+/-} (Hayashi et al., 2002). Las líneas de ratón *Dicer*^{fx/fx} (Cobb et al., 2005), *Dicer*^{+/-} y *Sox2Cre*^{+/-} (Hayashi et al., 2002) se mantuvieron en un fondo mixto. Para el genotipaje de los ratones se utilizaron los oligos previamente descritos para la detección del transgen *Sox2Cre* (Hayashi et al., 2002), y los detallados en la tabla 1 para la detección de los alelos del locus de *Dicer* silvestre (oligos Dicer B y C), mutante (oligos Dicer A y C) y condicional (fx, oligos Dicer B y C). La distinción entre los tres alelos se basa en el tamaño de DNA amplificado: 259 pb para el alelo silvestre, 309 pb para el alelo mutante y 390 pb para el alelo condicional. La reacción de PCR se llevó a cabo a una temperatura de anillamiento de 60°C en todos los casos.

Tabla 1. Oligos para el genotipaje de ratones en el locus de *Dicer*

Nombre del Oligo	Secuencia
Dicer A	5' AGTAATGTGAGCAATAGTCCCAG 3'
Dicer B	5' AGTGTAGCCTTAGCCATTTGC 3'
Dicer C	5' CTGGTGGCTTGAGGACAAGAC 3'

2 Hibridación *in situ* en embriones enteros

Embriones procedentes del cruce de machos y hembras *Dicer*^{+/-} se diseccionaron en medio M2 (Nagy et al., 2003) y se fijaron en paraformaldehído al 4% durante una noche. Al día siguiente se deshidrataron en series crecientes de metanol/PBS y se guardaron a -20°C en metanol al 100% hasta su uso.

La relación de sondas utilizadas y su procedencia se encuentra en la tabla 1. En cada caso se indica el nombre del gen, el tamaño de la sonda, laboratorio de procedencia y referencia de trabajos en los que se ha utilizado cuando los haya. La generación de las ribosondas marcadas con digoxigenina (DIG) se llevó a cabo siguiendo protocolos establecidos (Ariza-McNaughton and Krumlauf, 2002; Wilkinson, 1992). La hibridación *in situ* en cada caso se realizó sobre un total de 20 embriones escogidos al azar procedentes de un conjunto de mutantes, heterocigotos y silvestres.

La hibridación *in situ* se llevó a cabo siguiendo protocolos establecidos (Ariza-McNaughton and Krumlauf, 2002; Wilkinson, 1992). Los embriones se trataron con proteinasa K en función del estadio: los embriones en estadio E5.5 se trataron durante 1 minuto, los embriones en E6.5

durante 3 minutos y los embriones en estadio E7.5 durante 5 minutos. El revelado se llevó a cabo en solución de NBT/BCIP (Roche).

Tras el revelado de la hibridación *in situ* se tomaron fotografías de cada embrión al mismo aumento para a continuación ser genotipados por separado. Para el genotipaje se utilizó el protocolo descrito en (Martínez Barbera y col. 2000) y los oligos descritos en el apartado 1, con una temperatura de anillamiento de 58°C.

Tabla 2. Sondas utilizadas para caracterización del trofoblasto de embriones *Dicer*^{-/-}

Nombre del gen	Dominio de expresión	Tamaño	Laboratorio de origen	Referencia
<i>Cdx2</i>	TE	939 pb	Dra. J. Rossant	
<i>Esrrb</i>	TE	486 pb	Dra. J. Rossant	(Luo et al., 1997)
<i>Eomes</i>	TE, ExE, LP	1000 pb	Dra. E. Robertson	(Arnold et al., 2008)
<i>Ascl2</i>	EPC	2400 pb	Dr. F. Guillemot	(Guillemot et al., 1994)
<i>Rhox5</i>	ExE, EPC, exVE	650 pb	Dr. F. Guillemot	(Lin et al., 1994)
<i>Fgf4</i>	Epiblasto	950 pb	Dr. I. Mason	
<i>Fgfr2</i>	ExE, EPC	2000 pb	Dra. J. Rossant	(Ciruna and Rossant, 1999)
<i>Nodal</i>	Epiblasto	697 pb	Dra. E. Robertson	(Conlon et al., 1994)
<i>Spc1</i>	ExE	760 pb	Dr. D. Constam	(Roebroek et al., 1998)
<i>Spc4</i>	ExE	2940 pb	Dr. D. Constam	(Arnold et al., 2008)

Índice: TE-trofoectodermo; ExE-ectodermo extraembrionario; EPC-cono ectoplacentario; exVE-endodermo visceral extraembrionario; LP-línea primitiva

3 Análisis de proliferación con fosfo-histona 3 (PH3)

Se diseccionaron en M2 (Nagy et al., 2003) 22 embriones en estadio E6.5 procedentes del cruce de ratones *Dicer*^{+/-} y a continuación se cultivaron en un incubador a 37°C, 5%CO₂ durante 1 hora en medio DMEM (Gibco) con 50% de suero fetal bovino (Gibco) en presencia de Nocodazol a concentración 200nm (Sigma). Tras tres lavados en PBS se fijaron durante una noche en paraformaldehído al 4% a 4°C. Al día siguiente se lavaron en PBT (PBS-Tween-20 0.1%) y se incubaron en solución de bloqueo 1 (10% suero de cabra –Sigma-, 1% gelatina de pescado –Sigma-, 0.1% tritón –Sigma- en PBS) durante una hora a temperatura ambiente. A continuación se incubaron con anticuerpo anti-PH3 (Upstate) (1/500 en solución de bloqueo 1) durante una noche a 4°C. Al día siguiente se realizaron tres lavados cortos en solución de bloqueo 2 (1% gelatina de pescado, 0.1% tritón en PBS) y se incubaron los embriones con el anticuerpo secundario anti-rabbit-alexa568 (Invitrogen) (1/250 en solución de bloqueo 2) durante 4 horas a temperatura ambiente. Tras tres lavados cortos en PBT se dejaron los embriones durante una noche en gotas de medio de montaje “Vectashield with Dapi” (Vector Laboratories) en PBT.

El marcaje de PH3 de cada embrión se visualizó por microscopía confocal, seguida del genotipaje utilizando el mismo procedimiento que para las hibridaciones *in situ* (Martínez Barbera et al., 2000).

El recuento del número de células en división (PH3 positivas) se llevó a cabo con el programa ImageJ (Abramoff et al., 2004). Se analizaron todos los embriones mutantes (n=4) y 6 embriones silvestres. Se seleccionaron para cada embrión tres secciones con diferente nivel de proliferación para llevar a cabo el recuento, separadas entre sí por aproximadamente tres secciones (por ejemplo, recuento en las secciones 1, 4 y 8 de cada embrión). Los datos se expresaron como porcentaje de células en división y se hizo la media para embriones mutantes y silvestres, indicando en cada caso como error el Error de la Media. El contraste estadístico de las diferencias entre mutantes y silvestres se realizó con una prueba T-Student de dos colas ($p < 0.05$).

ANÁLISIS *IN VITRO*: FENOTIPO DE LAS CÉLULAS TS-*Dicer*^{fx/fx}

4 Derivación de las células TS-*Dicer*^{fx/fx}

4.1 Obtención de las líneas TSL1-*Dicer*^{fx/fx} y TSL4-*Dicer*^{fx/fx}

La preparación del medio TS y medio condicional, así como la inactivación de los fibroblastos embrionarios (MEFs) con mitomicina se encuentra detallado en (Himeno et al., 2008). La información referente al resto de reactivos básicos utilizados tanto para el cultivo como para la infección de células TS se encuentran en la tabla 3.

Tabla 3. Reactivos utilizados en el cultivo e infección de células TS

Reactivo	Concentración de uso	Referencia
0.05% Trypsin/1 mM EDTA	0.05%	Invitrogen
Human Recombinant FGF4	25 ng/ml	R & D Systems
Heparina	1 µg/ml	Sigma
RPMI 1640	-	Gibco
Human Factor X	9 µg/ml	Haematologic Technologies Incorporated
AdCre-GFP	70-100 MOI	University of Iowa Gene Transfer vector Core
Ad-GFP	70-100 MOI	University of Iowa Gene Transfer vector Core

La derivación de las células TS-*Dicer*^{fx/fx} a partir de blastocistos se llevó a cabo siguiendo protocolos establecidos (Himeno et al., 2008; Tanaka et al., 1998). Brevemente, ratones

Dicer^{fx/fx} se cruzaron para obtener un total de 22 blastocistos *Dicer*^{fx/fx} por lavado de útero a los 3.5 días post-fecundación (día 1 de derivación). Cada blastocisto se sembró sobre fibroblastos embrionarios (MEFs) inactivos. Los blastocistos se crecieron en medio TS suplementado con 25 ng/ml de FGF4 y 1 µg/ml de heparina hasta su salida de la zona pelúcida y adhesión a la placa formando un sobre-crecimiento celular conocido como “outgrowth”. En el día 4 de la derivación 19 de 22 blastocistos habían formado un “outgrowth”. Estos 19 “outgrowth” se disgregaron con tripsina 0.05% y las células resultantes de cada uno se sembraron de nuevo sobre MEFs inactivos con medio TS suplementado con 70% de medio condicionado, 25 ng/ml de FGF4 y 1 µg/ml de heparina. Las células se mantuvieron de este modo renovando el medio cada día durante los dos días siguientes y cada dos días de ahí en adelante. A los 11 días comenzaron a verse en 8 de las 19 placas de partida colonias de células TS, que se distinguen por tener una morfología de colonia epitelial característica así como por el tamaño y forma de las células que forman las colonias, muy diferentes de las células ES y XEN también derivadas de blastocistos. Se continuó cultivando del mismo modo estas 8 líneas pasándolas a placas de mayor tamaño en función del grado de confluencia alcanzado por las colonias de células TS en cada caso. Cuando las células alcanzaron el pase 3 ó 4 se comenzaron a cultivar con el mismo medio en ausencia de MEFs inactivos. A los dos meses de comenzar la derivación únicamente 2 líneas, TSL1-*Dicer*^{fx/fx} y TSL4-*Dicer*^{fx/fx} pudieron mantenerse en condiciones de crecimiento y auto-renovación normales durante 7 y 9 pases respectivamente.

Las líneas TSL1-*Dicer*^{fx/fx} y TSL4-*Dicer*^{fx/fx} se mantuvieron durante dos meses más en cultivo (un total de cuatro meses) siguiendo sucesivos pases y rondas de congelación/descongelación hasta que finalmente se consideraron como líneas estables al alcanzar alrededor de 16 ó 17 pases. Únicamente entonces se consideraron adecuadas para la realización de experimentos.

4.2 Caracterización de las líneas TSL1-*Dicer*^{fx/fx} y TSL4-*Dicer*^{fx/fx}

Previo al comienzo de los experimentos con las líneas TSL1-*Dicer*^{fx/fx} y TSL4-*Dicer*^{fx/fx} se confirmó la calidad de las mismas en cuanto a proliferación, capacidad de auto-renovación y diferenciación. Ambas líneas forman las colonias epiteliales características de células TS y son capaces de diferenciarse a células gigantes de trofoblasto cuando se retira el suplemento de Fgf4 o el medio condicionado del medio de cultivo (Fig. 1A-B’).

Dado que las células gigantes de trofoblasto se caracterizan por endo-reduplicar su DNA, el nivel de diferenciación en este tipo celular en condiciones de cultivo normales para ambas líneas se midió mediante un análisis del contenido en DNA (Tanaka et al., 1998). Brevemente, las células se disgregaron con tripsina se lavaron y precipitaron. Las células precipitadas se fijaron con etanol al 70% frío seguido de lavados con PBS y finalmente se resuspendieron en solución de tinción de yoduro de propidio (50 µg/ml de yoduro de propidio –Sigma-, 0.05% de NP-40 –Fluka-, y 1 mg/ml de RNasa A –Sigma-, en PBS), donde se mantuvieron en oscuridad durante 10 minutos a temperatura ambiente y 20 minutos en hielo. A continuación se analizó el contenido en DNA mediante citometría de flujo. El resultado (Fig. 1C, D) muestra un nivel de

diferenciación (medido por un contenido en DNA mayor de 4n) del 20% para la línea TSL1-*Dicer*^{fx/fx} y del 22% para la línea TSL4-*Dicer*^{fx/fx}, en ambos casos comparable al nivel obtenido para la línea previamente descrita TS-B1 (Mak et al., 2002) mantenida en cultivo en el laboratorio del Dr. T. Rodríguez (24% de diferenciación, datos no mostrados).

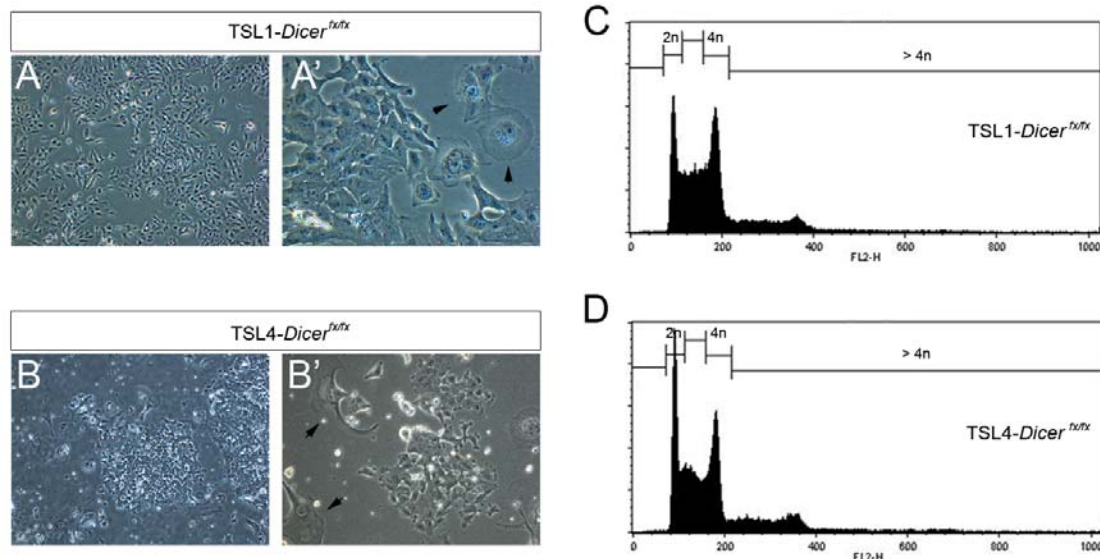


Figura 1. Caracterización de las líneas TSL1-*Dicer*^{fx/fx} y TSL4-*Dicer*^{fx/fx}. Las células TSL1-*Dicer*^{fx/fx} y TSL4-*Dicer*^{fx/fx} forman colonias de morfología epitelial características de células TS (A, B) y son capaces de diferenciarse en células gigantes de trofoblasto tras la retirada del Fgf4 o del medio condicionado (indicadas con flechas en A' y B'). El perfil de ciclo celular es similar al obtenido para células TS previamente descritas (Tanaka et al., 1998), y en ambos casos obtenemos un porcentaje de células (20% en el caso de TSL1-*Dicer*^{fx/fx} y 22% para TSL4-*Dicer*^{fx/fx}) con un contenido en DNA >4n, correspondiente a células poliploides diferenciadas (células gigantes de trofoblasto) (C,D). Este porcentaje se encuentra dentro de lo esperado para un cultivo de una línea estable de células TS.

5 Deleción de *Dicer* in vitro: sistema de infección con adenovirus

Las células TS-*Dicer*^{fx/fx} se infectaron con el adenovirus AdCre-GFP para la deleción de *Dicer* o con el adenovirus AdGFP como control negativo. Ambos adenovirus se obtuvieron de la Universidad de Iowa Gene Transfer Vector Core (<http://www.uiowa.edu/~gene/>) (Tabla 3).

En la figura 2 se muestra un esquema del protocolo llevado a cabo en los experimentos de infección con adenovirus. Las células TS-*Dicer*^{fx/fx} se sembraron 24h antes de la infección de forma que alcanzasen un 60-70% de confluencia en el momento de la infección (35×10^4 células por placa en placas de 35mm; 1×10^6 células por placa en placas de 10cm). El día de la infección se sustituyó el medio completo por medio RPMI 1640 (sin suero y sin antibióticos), donde se mantuvieron las células durante 4h previas a la infección. En el momento de la infección se diluyó el virus correspondiente (en una concentración de 70-100 MOI) en medio RPMI 1640 conteniendo 9µg/ml de Human Factor X (Tabla 3). La infección se llevó a cabo en 1ml para placas de 35mm y el 4.5ml para placas de 10cm. Se reemplazó el medio de las

células por medio conteniendo el virus y se incubaron durante 4 horas a 37°C, 5%CO₂. Tras las cuatro horas de incubación se retiró el medio con virus y se añadió medio completo (70% medio condicional, 25 ng/ml de FGF4, 1µg/ml de heparina en medio TS). Transcurridas 24h de la infección se seleccionaron por citometría de flujo las células GFP positivas para la infección con cada virus y se sembraron en igual densidad en tres placas independientes en cada caso. Cada una de estas tres placas se utilizó para el recuento del número de células y extracción de RNA a los días 3, 5 y 7 tras la infección. La eficiencia de la infección siguiendo este protocolo fue del 80-90%. En el caso de las células infectadas con AdCre-GFP, cada muestra recogida se genotipó en paralelo (ver apartado 1 para genotipaje) para confirmar la delección de *Dicer*.

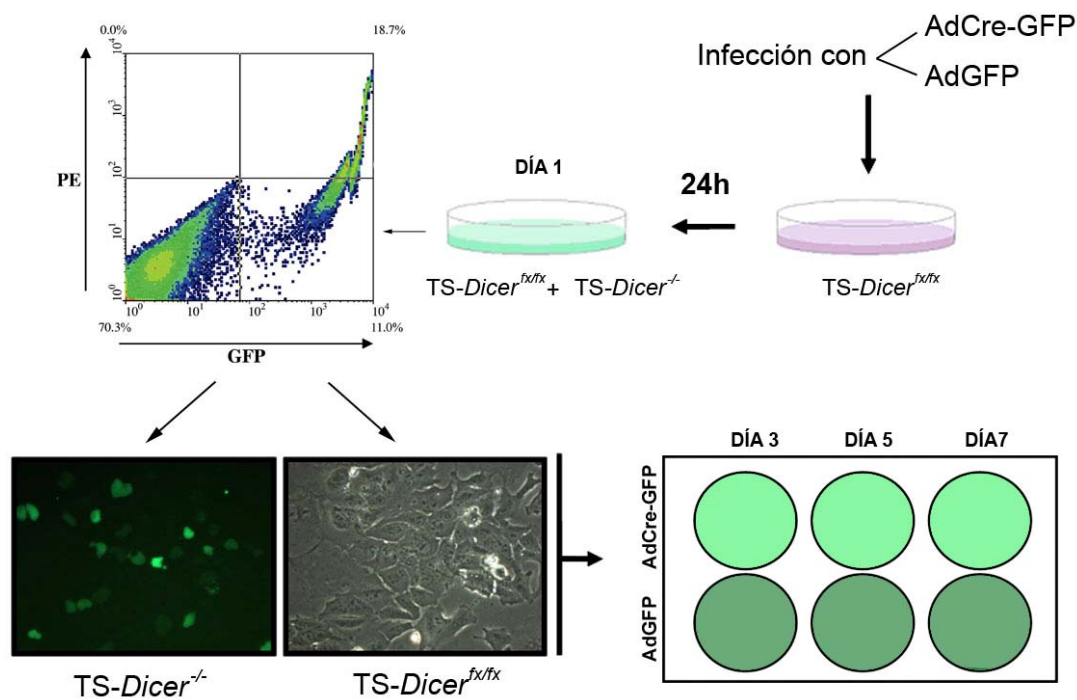


Figura 2. Protocolo llevado a cabo para la infección de las células TS-Dicer^{fx/fx}. Las células TS-Dicer^{fx/fx} se infectaron con el adenovirus AdCre-GFP (generación de mutantes) o AdGFP (control negativo). A las 24h de la infección (Día 1) se seleccionaron por citometría de flujo en ambos casos las células GFP positivas, que se sembraron en igual densidad y en tres placas en cada caso para el análisis de las muestras (recuento del número de células y extracción de RNA) a los 3, 5 y 7 días tras la infección.

Los resultados obtenidos de las primeras infecciones realizadas mostraron un comportamiento idéntico en cuanto a diferenciación y crecimiento para las líneas TSL1-Dicer^{fx/fx} y TSL4-Dicer^{fx/fx} infectadas (datos no mostrados), por lo que en sucesivos experimentos se utilizó únicamente la línea TSL4-Dicer^{fx/fx}, a la que nos referimos como TS-Dicer^{fx/fx}.

6 Análisis de las células TS-Dicer^{-/-}

6.1 Recuento del número de células y extracción de RNA

Transcurridos 3, 5 y 7 días desde la infección se disgregaron las células infectadas con AdCre-GFP y con AdGFP con tripsina y se contaron utilizando una cámara de Neubauer. A continuación se precipitaron y resuspendieron en 800µl de Trizol (Invitrogen) conteniendo 0.3µl de tRNA (10mg/ml) para llevar a cabo la extracción de RNA de acuerdo al protocolo indicado por el fabricante (Invitrogen). El RNA extraído de esta forma se trató con Turbo DNasa (Ambion) de acuerdo a las instrucciones del fabricante durante 45 minutos a 37°C y se limpió a continuación siguiendo el protocolo RNA Cleanup del kit de extracción de RNA RNeasy MiniKit (Qiagen).

6.2 PCR cuantitativa

Las muestras de RNA extraídas como se indica más arriba se utilizaron para la generación de cDNA. Para ello se añadió 1.5µg de RNA y 1µl de Nanomer Primers (Invitrogen) a un volumen total de 10µl de agua y se mantuvo 5 minutos a 65°C seguido de 1 minuto en hielo. A continuación se añadió una mezcla de 2µl de dNTPs (Invitrogen), 2µl de DTT 0.1M (Invitrogen), 1µl de inhibidor de RNasa (RNase inhibitor, Roche) y 1µl de enzima *SuperScript*® III Reverse Transcriptase (Invitrogen) en un volumen total de 10µl de agua y se incubó 5 minutos a 25°C, 1 hora 30 minutos a 42°C y por último 15 minutos a 72°C. Para el control negativo de transcripción se utilizó la misma mezcla de reacción sin enzima *SuperScript*® III RT. El cDNA así generado se diluyó 1/10 para su uso en la reacción de PCR cuantitativa.

Para la reacción de PCR cuantitativa se utilizaron 2µl de cDNA, 0.9µl de cada oligo (a concentración 10µM), 15µl de enzima SYBR Green (Qiagen) y 11.2µl de agua. Los oligos utilizados se encuentran detallados en la tabla 4. En cada caso se incluyó una muestra con el control negativo de retrotranscripción. Para cada pareja de oligos se realizó una curva de calibración previa con diluciones crecientes de cDNA para comprobar una correcta variación de la señal en función de la concentración de cDNA. Se utilizaron como genes de normalización los genes *Hmbs* (Hidroximetilbilano Sintasa) y *Ywhaz* (proteína de activación de tirosina 3-monooxigenasa/triptófano 5-monooxigenasa, polipéptido z) (Vandesompele et al., 2002).

La PCR cuantitativa se llevó a cabo en la máquina Opticon DNA utilizando un programa convencional de 40 ciclos a 59°C, incluyendo en cada caso una curva de disociación. Para el análisis inicial de los datos de amplificación se utilizó el programa MJ Opticon Monitor Analysis Software 3.1 (MJ Research).

Los resultados de PCR cuantitativa mostrados son los resultantes de la media de tres experimentos independientes. En todos los casos el Cambio positivo o negativo del nivel de mRNA en células infectadas con AdCre-GFP respecto a las infectadas con AdGFP se calculó como $2^{-\Delta Ct}$, siendo ΔCt la diferencia entre los niveles de expresión en ambas muestras.

Tabla 4. Oligos utilizados en el análisis por PCR cuantitativa.

Gen	Secuencia de los oligos
<i>Cdx2</i>	F: 5' TTCACCTTCTCAGAGACACAGTTCAT 3' R: 5' GAGTTAACCTGTCATTTTCTGAAGCC 3'
<i>Esrrb</i>	F: 5' GGACACACTGCTTTGAAGCA 3' R: 5' ACAGATGTCTCTCATCTGGC 3'
<i>Eomes</i>	F: 5' TTCACCTTCTCAGAGACACAGTTCAT 3' R: 5' GAGTTAACCTGTCATTTTCTGAAGCC 3'
<i>Fgfr2</i>	F: 5' GAGGAATACTTGGATCTACC 5' R: 5' CTGGTGCTGTCCTGTTTGGG 5'
<i>Stra13</i>	F: 5' GGTGAGCAGACTACTCCATT 3' R: 5' GTGCCCCACATATTCCCCAC 3'
<i>Hand1</i>	F: 5' TGCACAAGCAGGTGACCCCG 3' R: 5' CCCTTTAATCCTCTTCTCGCCG 3'
<i>Pl1</i>	F: 5' TACCCTGCTTGGTCTGGACT 3' R: 5' GGGCACTCAACATTGTTCT 3'
<i>Pai-1</i>	F: 5' CTCCACAGCCTTTGTCATCTCA 3' R: 5' CCGAACCACAAAGAGAAAGGA 5'
<i>Dicer1</i>	R: 5' TATCGCCTTCACTGCCTTCT 3' F: 5' GCAACCTTTTGCAGTTCACA 3'
<i>Hmb</i>	F: 5' ACTGGTGGAGTCTGGAGTCTAGATGGC 3' R: 5' GCCAGGCTGATGCCAGGTT 3'
<i>Ywhaz</i>	F: 5' CGTTGTAGGAGCCCGTAGGTCAT 3' R: 5' TCTGGTTGCGAAGCATTGGG 3'

7 Variación en mRNA en células TS-*Dicer*^{-/-}: array de mRNA

El análisis de variación en expresión génica a nivel de RNA mensajero tras la delección de *Dicer* se llevó a cabo mediante la realización de un array en células infectadas TS-*Dicer*^{-/-} con AdCre-GFP frente a células infectadas con AdGFP. Para ello se hicieron infecciones siguiendo el protocolo indicado en el apartado 5 y se extrajo el RNA como se detalla en el apartado 6.1. Se analizó la expresión a días 4 y 6 después de la infección. Para ello se recogieron a estos dos tiempos tres muestras independientes de células TS-*Dicer*^{fx/fx} infectadas con AdCre-GFP y tres muestras infectadas con AdGFP en las que los experimentos de infección con cada virus se realizaron simultáneamente. El RNA extraído de estas muestras se envió al laboratorio UCL

Genomics (Institute of Child Health, UCL), para su procesamiento e hibridación en el array Exon Expression Array 1.0ST de la plataforma Affimetrix.

Los datos obtenidos del array fueron procesados estadísticamente y transformados en valores de cambio positivo o negativo por la Dra. M. Leleu (personal de apoyo informático del Clinical Science Centre-MRC).

7.1 Análisis de Gene Ontology (GO)

El análisis de Gene Ontology se llevó a cabo sobre aquellos genes cuyo aumento en expresión a día 4 tras la infección fuese superior a un cambio de 1.3 veces ($p < 0.05$). La elección de un límite de cambio bajo se hizo teniendo en cuenta que la represión que ejercen los miRNAs a nivel de mRNA va a ser siempre inferior a la represión a nivel de proteína, y por tanto el efecto sobre la cantidad de mRNA será en muchos casos más débil (Baek et al., 2008; Selbach et al., 2008).

Para el análisis de Gene Ontology se utilizó el programa FatiGO, dentro de la plataforma BABELOMICS (www.fatigo.org) (Al-Shahrour et al., 2006; Al-Shahrour et al., 2005). Se analizaron los términos de Proceso Biológico (niveles 3 a 9) sobre-representados en la lista de genes cuya expresión aumenta a los 4 días de la infección frente al total de genes en el genoma. Se aplicó el test estadístico exacto de Fisher (p -valor corregido < 0.05) para determinar qué términos se encuentran sobre-representados en nuestra lista respecto al resto del genoma.

Los términos sobre-representados en un nivel más alto de GO (nivel 3) se clasificaron en distintas categorías en función de los términos de categorías inferiores que se incluyen dentro de cada nivel superior.

EXPRESIÓN DE miRNAs EN CÉLULAS TS Y EMBRIÓN TEMPRANO

8 Análisis de expresión de miRNAs: array por PCR Cuantitativa

Las líneas MEF-CD1 y MEF-T utilizadas para el array de expresión de miRNAs fueron derivadas en los laboratorios del Dr. J. Gil y la Dra. V. Yu respectivamente. La línea de ES-m5 se derivó en el laboratorio del Dr. T. Rodríguez. Las líneas ES-E14 y TS-B1 se han descrito previamente (Mak et al., 2002; Sawai et al., 1991) y fueron cedidas por la Srta. Z. Webster y el Dr. N. Brockdorff respectivamente. Para las muestras de embriones se diseccionaron camadas procedentes de cruces de ratones CD1.

Para la extracción del RNA de células éstas se crecieron en todos los casos hasta alcanzar una confluencia del 70-80%. Llegado ese punto se disgregaron con tripsina 0.05%, se precipitaron y se lavó el precipitado para a continuación comenzar con el proceso de extracción de RNA. Los embriones se diseccionaron en medio M2 (Nagy et al., 2003), se lavaron brevemente en PBS y se congelaron en nitrógeno líquido hasta acumular un total de

aproximadamente 200 embriones en estadio E5.5, 100 en estadio E6.5, 20 en estadio E7.5 y 3 en estadio E8.5. Una vez alcanzado el número de embriones necesario se extrajo el RNA. La extracción de RNA se llevó a cabo en todos los casos con el kit *mirVana*[™] miRNA Isolation Kit (Ambion) siguiendo el protocolo de extracción de RNA total indicado por el fabricante.

Las muestras de RNA extraídas de esta forma se enviaron a la compañía Geneservice (www.geneservice.co.uk) para su procesamiento y análisis mediante el array de expresión de miRNAs por PCR Cuantitativa “TaqMan[®] Mouse MicroRNA Assay Set v1.0” (Applied Biosystems). Este array cubre la expresión de 312 miRNAs bien documentados en ratón. La información referente a los miRNAs analizados en este array se encuentra en la página <https://products.appliedbiosystems.com/ab/en/US/adirect/ab?cmd=catNavigate2&catID=60180>. El fundamento del sistema de amplificación se encuentra esquematizado en la figura 3.

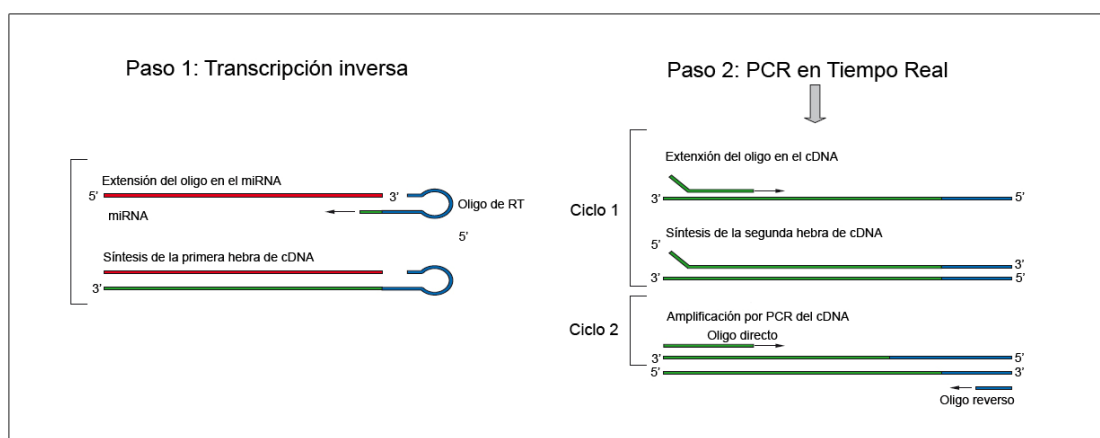


Figura 3. Proceso de amplificación de miRNAs. El protocolo para la amplificación de miRNAs por PCR cuantitativa consta de dos pasos. En un primer paso un oligo especialmente diseñado se acopla al extremo 3' del miRNA y se elonga en un proceso de transcripción inversa. En un segundo paso se lleva a cabo una PCR en tiempo real utilizando como molde el cDNA previamente generado. Imagen modificada del Protocolo para TaqMan[®] MicroRNA Assays (Applied Biosystems).

Como control interno se utilizaron snoRNAs (RNA pequeño nuclear), puesto que sufren el mismo proceso de transcripción inversa que los miRNAs. En nuestro caso los datos se normalizaron frente al sno135. El resultado final es el resultante de la media de tres experimentos de amplificación.

8.1 Análisis de datos del array de miRNAs

El agrupamiento jerárquico de los perfiles de expresión de las distintas muestras se llevó a cabo con el programa MultiExperimental Viewer Version 4.1 (MeV v4.1). Para ello se utilizaron valores de Ct normalizados frente al control interno sno135. Se utilizó el coeficiente de correlación como medida de similitud y ligamiento completo como método de agrupamiento.

Las distribuciones de frecuencias para los niveles de expresión de cada muestra se generaron utilizando el programa GraphPad Prism 5. Las distribuciones se basan en los valores de Ct obtenidos para cada muestra analizada, con un ancho de bin de 2 en un rango de -4 a 18.

En los casos en que los valores de expresión de miRNAs están indicados como Expresión Relativa, ésta se calculó como 2^{-Ct} . En los casos en que se establece una comparación entre los niveles de expresión en dos tipos celulares (por ejemplo TS y ES) los valores de Cambio se calcularon como $2^{-\Delta Ct}$, siendo ΔCt la diferencia entre los valores de expresión de ambos tipos celulares.

8.2 Patrones de expresión de miRNAs

La detección del patrón de expresión de los miRNAs se llevó a cabo utilizando sondas de LNA obtenidas sin marcar de la compañía Exiqon. La secuencia de cada LNA, su temperatura de anillamiento (T_m) y la temperatura a la que se llevó a cabo la hibridación *in situ* (T Hibridación) se encuentran detalladas en la tabla 5.

Tabla 5. Sondas LNA para la detección de miRNAs por hibridación *in situ*

miRNA	Secuencia LNA	T_m (°C)	T Hibridación (°C)
miR-467b	5' CATATACATGCAGGCACTTAC 3'	71	50
miR-669a	5' ACATGAACATGCACACACAACT 3'	76	55
miR-291-3p	5' GCACACAAAGTGAAGCACTTT 3'	75	54
miR-291-5p	5' AGAGAGGGCCTCCACTTTGATG 3'	77	56
miR-295	5' AGACTCAAAAGTAGTAGCACTTT 3'	70	49

Los embriones utilizados proceden de cruces de BL6xCBA que se diseccionaron, fijaron y deshidrataron como se indica en apartados anteriores (ver apartado 2).

Las sondas LNA se marcaron en el extremo 3' utilizando el kit "DIG Oligonucleotide 3'-End Labeling Kit" (Roche), siguiendo las recomendaciones del fabricante.

La hibridación *in situ* con las sondas LNA se llevó a cabo utilizando protocolos previamente descritos (Kloosterman et al., 2006) con las modificaciones que se describen a continuación. Tras la rehidratación de los embriones éstos se blanquearon durante 1h con peróxido de hidrógeno al 6% a temperatura ambiente. Tras lavados en PBT (PBS-Tween-20 0.1%) los embriones se trataron con proteinasa K (10µg/ml en PBT) en función del estadio: 1 minuto para E5.5, 3 minutos para E6.5, 5 minutos para E7.5, seguido de tratamiento con glicina (2mg/ml en PBT) durante 5 minutos. Los embriones se refijaron durante 20 minutos en 0.2% glutaraldehído/4% paraformaldehído en PBT. Previa a la hibridación con la sonda se llevó a cabo una prehibridación durante al menos 1 hora 30 minutos en solución de hibridación (Formamida 50%, SSC 5x, Tween-20 0.1%, Heparina 50µg/ml, tRNA µg/ml, Acido cítrico hasta

pH6.0). La temperatura de pre-hibridación e hibridación fue de 20-22°C inferior a la T_m indicada para cada sonda LNA (ver tabla 5). La hibridación se llevó a cabo durante una noche en 200µl de solución de hibridación conteniendo 1.5µl de sonda marcada como se indica más arriba. Los siguientes pasos del protocolo se llevaron a cabo de acuerdo con lo descrito en (Kloosterman et al., 2006), revelando el cuarto día en solución de NBT/BCIP (Roche).

Los embriones se seccionaron tras la hibridación *in situ*. Para ello cada embrión se embebió en 30% sacarosa durante una noche a 4°C y se incluyó en gelatina 7.5%- sacarosa 15% en PBS. Los bloques generados de este modo se congelaron en isopentano para llevar a cabo las secciones en criostato.

La sonda para la detección de la expresión del gen *Sfmbt2* se generó por amplificación por PCR de cDNA de células TS-B1 producido por retrotranscripción como se indica en el apartado 6.2. La reacción de amplificación por PCR se llevó a cabo a una temperatura de 60°C utilizando los oligos: Sfmbt2-F: 5' CTGATAAACCCACGATGCTGT 3', Sfmbt2-R: 5' GCTGCATTGGTAATGAGGGTA 3' (tamaño amplificado de 703 pb). El producto de la amplificación se clonó en el vector pGEM-T Easy Vector (Promega) para ser utilizado como molde para transcripción en la generación de la sonda. La transcripción de la sonda e hibridación *in situ* se llevó a cabo utilizando protocolos establecidos tal y como se indica en el apartado 2.

miRNAs IMPLICADOS EN LA FUNCIÓN DE CÉLULAS TS

9 miRNAs asociados a la variación génica de las células TS-*Dicer*^{-/-}

En la figura 4 se representa un esquema de los pasos seguidos para la obtención de los datos finales, procedentes de la asociación de los datos obtenidos de los arrays de mRNA y miRNA con los miRNAs predichos.

El análisis de los miRNAs implicados en el aumento de la expresión génica en las células TS-*Dicer*^{-/-} se llevó a cabo utilizando los archivos correspondientes a los sitios conservados y no conservados predichos presentes en la base de datos TargetScan 4.1 (www.targetscan.org). En esta base de datos se encuentran todos los sitios predichos en regiones 3'UTR para todos los miRNAs incluyendo toda la información referente a unión de un determinado miRNA a su mensajero diana. En función de la fiabilidad de la predicción teniendo en cuenta distintos parámetros, a cada sitio de unión para un determinado miRNA se le adjudica un valor denominado "context score", que se considera mejor a medida que aumenta en sentido negativo. El algoritmo utilizado para la predicción se encuentra detallado en (Grimson et al., 2007), y de acuerdo a lo expuesto en dicho trabajo se puede considerar que un sitio predicho puede ser real a partir de un valor de "context score" inferior a -0.1. Basándonos en esta predicción utilizamos para el análisis todos los sitios para miRNAs con un valor de "context

score" igual o inferior a -0.1 (Fig. 4). Estos datos correspondientes a sitios predichos se cruzaron con los datos obtenidos del array de mRNA en células TS-*Dicer*^{-/-} a día 4 tras la delección, utilizando únicamente aquellos genes con expresión aumentada con un cambio superior a 1.3 veces ($p < 0.05$). De este modo se obtuvo una lista de sitios predichos para miRNAs en el set de genes con expresión aumentada. El nivel de expresión para cada miRNA en condiciones normales (células TS-*Dicer*^{+/+}) se obtuvo de los datos procedentes del array por PCR cuantitativa previamente realizado.

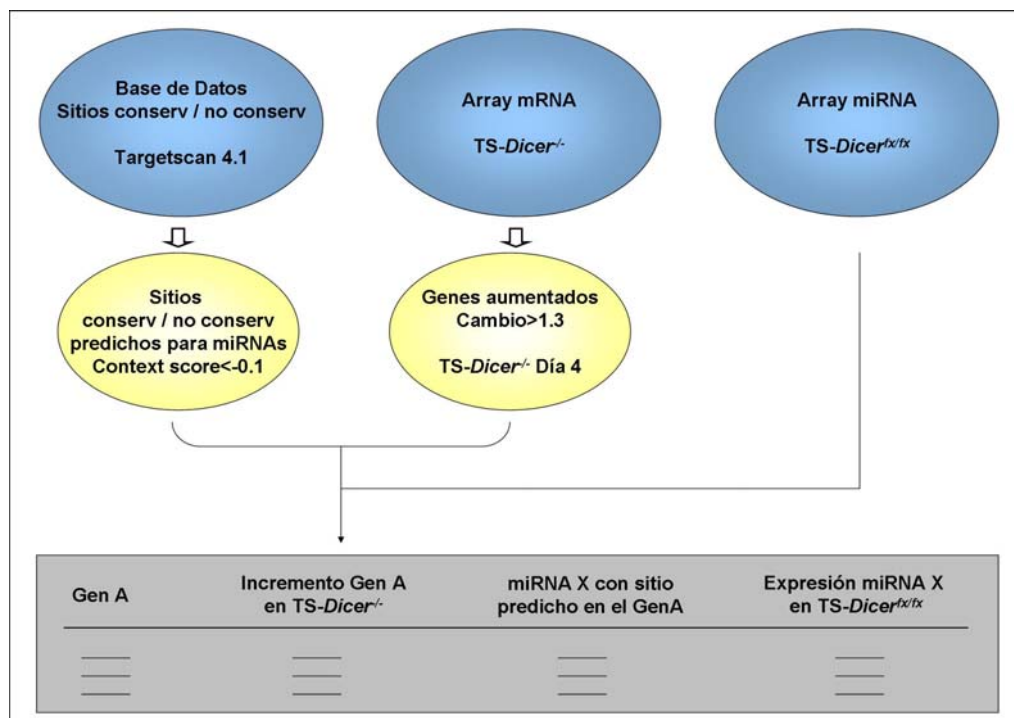


Figura 4. Esquema de asociación de datos. Los datos correspondientes a los genes aumentados en TS a los 4 días de la delección de Dicer (Cambio > 1.3 veces, $p < 0.05$) se cruzaron con los datos correspondientes a sitios conservados o no conservados para miRNAs predichos obtenidos de la base de datos del programa TargetScan 4.1 (www.targetscan.org), utilizando sólo sitios predichos con un valor de "context score" < -0.1. La expresión de cada miRNA en células TS-*Dicer*^{+/+} se obtuvo a partir del array previamente realizado por PCR cuantitativa.

La asociación de todos los datos se llevó a cabo utilizando el programa Excel mediante una macro especialmente diseñada que permite la búsqueda y asociación de datos procedentes de diferentes hojas de cálculo.

El análisis estadístico de los sitios significativos para cada miRNA se realizó estableciendo una relación entre los sitios observados en los genes aumentados en las células TS-*Dicer*^{-/-} y los esperados por azar en función del total de sitios para cada miRNA en todos los genes de la base de datos, utilizando un test estadístico Chi-Cuadrado con corrección de Yates ($p < 0.05$).

Resultados

ANÁLISIS *IN VIVO* DEL FENOTIPO DE FALTA DE FUNCIÓN DE DICER EN EL TROFOBlasto

En la última década diferentes grupos de investigación se han aproximado al estudio de la función de los miRNAs durante el desarrollo embrionario mediante la generación de organismos o tejidos carentes de estos reguladores. Para ello una de las estrategias más comúnmente usadas ha sido la interferencia con su maquinaria de producción por medio de la delección de genes implicados en la ruta de biosíntesis. Entre estos genes se encuentra *Dicer*, encargado de la maduración en el citoplasma de los miRNAs a partir de su precursor (pre-miRNA). De este modo aquellas mutaciones que interfieran con la función de *Dicer* tendrán como resultado el impedimento en la maduración de los miRNAs, y el fenotipo resultante vendrá dado fundamentalmente por el efecto de falta de función de estos reguladores.

Se han generado distintas cepas de ratón mutantes para *Dicer*, en su mayoría portadoras de alelos condicionales para la posterior generación del mutante completo o condicional en tejidos específicos (Bernstein et al., 2003; Cobb et al., 2005; Harfe et al., 2005; Mudhasani et al., 2008). En todas ellas la delección de *Dicer* en el embrión completo trae consigo una parada de crecimiento y muerte a los 7.5 días de desarrollo. Este fenotipo temprano fue descrito por primera vez por Bernstein y col. en 2003, sin embargo hasta la fecha no ha sido estudiado en detalle.

En este primer apartado describiremos el fenotipo del mutante completo de *Dicer* generado a partir de la línea condicional generada por Cobb y colaboradores en 2005, centrando nuestro estudio en los efectos sobre el desarrollo de uno de los principales linajes especificados durante las primeras etapas del desarrollo embrionario: el trofoblasto extraembrionario que dará lugar a la placenta.

1 Los embriones *Dicer*^{-/-} son morfológicamente anormales y mueren en gastrulación

Con el objetivo de obtener embriones mutantes para la proteína Dicer utilizamos una línea transgénica condicional previamente descrita (*Dicer*^{fx/fx}, (Cobb et al., 2005)) en la que dos sitios loxP se encuentran flanqueando uno de los dominios RNasaIII, esencial para la función de la proteína. Del cruce de machos homocigotos para el alelo condicional de *Dicer* con hembras portadoras del alelo *Sox2-Cre* (Hayashi et al., 2002), heredado de forma materna, se obtuvieron los ratones *Dicer*^{+/-}.

Los ratones *Dicer*^{+/-} son indistinguibles de los *Dicer*^{+/+}, y se cruzaron para analizar la descendencia *Dicer*^{-/-}. A los 5.5 días de desarrollo (estadio E5.5) un 78% de los embriones *Dicer*^{-/-} tienen un tamaño inferior a los silvestres (n=14/18) pero son morfológicamente normales (Fig. 1A, A'). Sin embargo a los 6.5 días de desarrollo (E6.5) comienzan a hacerse aparentes diversas malformaciones en los embriones mutantes (Fig. 1B, B'). Éstos son de tamaño inferior

a los silvestres y el epiblasto adquiere una forma redondeada, con la unión entre el ectodermo extraembrionario y el epiblasto marcada por una constricción pronunciada en un 30% de los casos ($n=8/26$). Los embriones *Dicer*^{-/-} a los 7.5 días de desarrollo (estadio E7.5) son de tamaño muy inferior a los silvestres y presentan evidentes defectos morfológicos que varían entre individuos (Fig. 1C, C'). En un 45% de los casos se observó una reducción marcada del ectodermo extraembrionario ($n=5/11$) sin verse aparentemente afectada la región de diferenciación del cono ectoplacentario. Un 29% ($n=16/54$) de los embriones recogidos en este estadio eran mutantes, lo que indica una supervivencia hasta los 7.5 días de desarrollo con proporciones mendelianas. Más allá de estadio E7.5 no se recuperaron embriones *Dicer*^{-/-} y en su lugar se obtuvieron deciduas vacías junto a los embriones heterocigotos o silvestres de la misma camada.

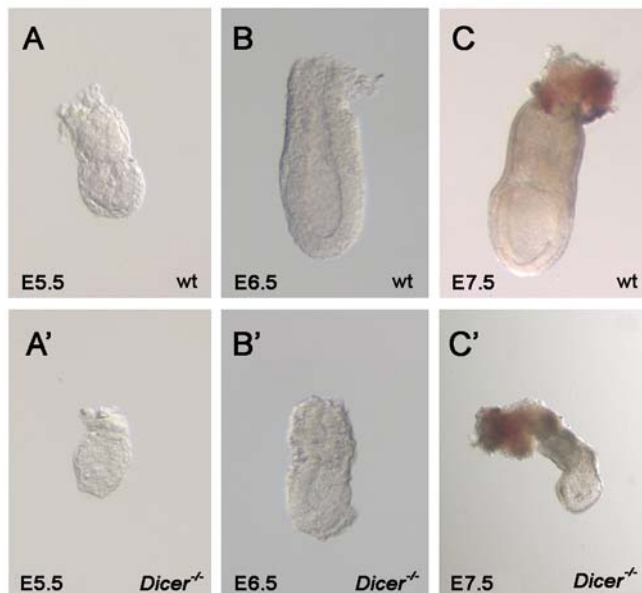


Figura 1. Los embriones *Dicer*^{-/-} son morfológicamente anormales. Los embriones *Dicer*^{-/-} son de tamaño inferior a los silvestres desde estadio E5.5 (**A**, **A'**). En estadio E6.5 comienzan a observarse defectos morfológicos en el epiblasto, caracterizados en muchos casos por un epiblasto redondeado y una constricción marcada en la región límite entre el epiblasto y el trofoblasto (**B**, **B'**). En estadio E7.5 el tamaño de los mutantes es muy inferior al de los silvestres, el epiblasto tiene forma redondeada y el ectodermo extraembrionario se encuentra reducido (**C**, **C'**).

Así pues, los embriones *Dicer*^{-/-} tienen un tamaño reducido respecto a los silvestres y presentan defectos morfológicos aparentes en el epiblasto desde el estadio E6.5 y más acusados en el estadio E7.5, en el que también se aprecia una reducción en el ectodermo extraembrionario. Finalmente, los embriones *Dicer*^{-/-} no sobreviven más allá de los 7.5 días de desarrollo.

2 Defectos en el trofoblasto de embriones *Dicer*^{-/-}

2.1 Falta de expresión de genes de multipotencia

En etapas tempranas de post-implantación, en el trofoblasto adyacente al epiblasto se encuentra una población de células madre con potencial para diferenciarse en todos los linajes de la placenta, las células TS (Tanaka et al., 1998). El mantenimiento de esta población depende de la expresión de ciertos genes. Entre ellos son de especial relevancia los factores

de transcripción *Cdx2* y *Eomes* y el receptor de estrógenos *Esrrb*, sin los cuales el embrión es incapaz de mantener las células TS y por tanto de especificar correctamente el resto de linajes del trofoblasto y la placenta (Luo et al., 1997; Russ et al., 2000; Strumpf et al., 2005). El análisis por hibridación *in situ* en embriones enteros en estadio E6.5 revela una pérdida de expresión de *Cdx2* (n=4/4), *Esrrb* (n=6/7) y *Eomes* (n=3/5) en los embriones *Dicer*^{-/-} (Fig. 2A-C').

Entre los genes expresados en las poblaciones diferenciadas del trofoblasto se encuentran los factores de transcripción *Ascl2*, restringido al cono ectoplacentario y *Rhox5*, que se expresa tanto en el ectodermo extraembrionario como en el cono ectoplacentario y en endodermo visceral extraembrionario (Guillemot et al., 1994; Lin et al., 1994). Al contrario de lo que sucede con los genes específicos de células TS, *Ascl2* y *Rhox5* presentan expresión normal en todos los embriones mutantes a los 6.5 días de desarrollo (Fig. 2D-E'; *Ascl2*, n=4/4; *Rhox5*, n=3/3).

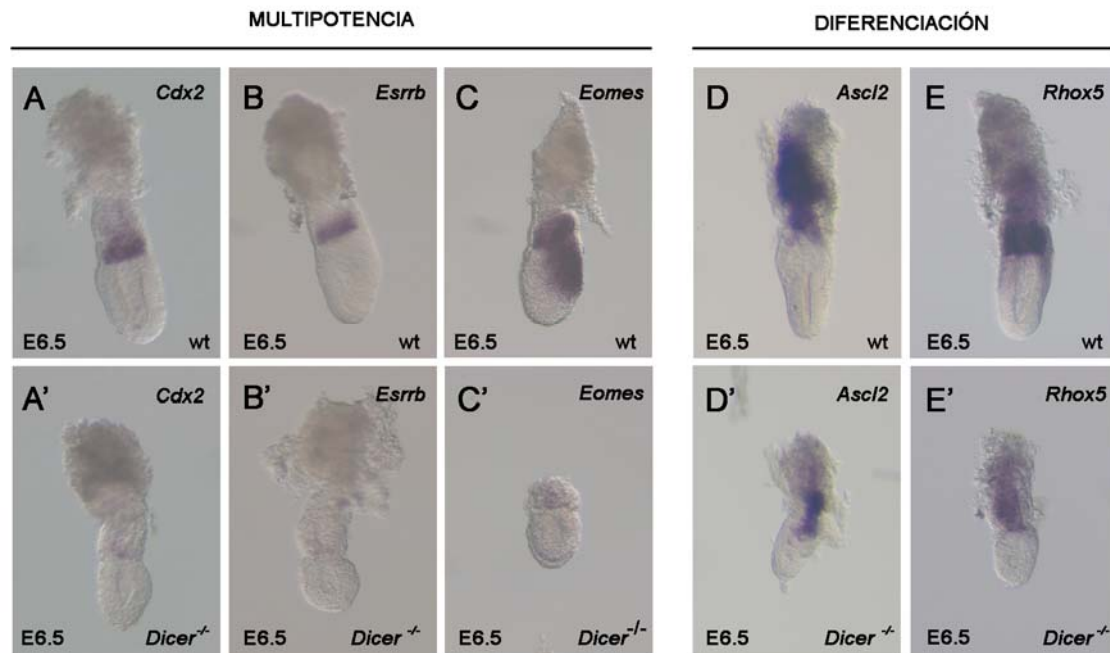


Figura 2. Los embriones *Dicer*^{-/-} tienen defectos específicos de expresión de genes de multipotencia en estadio E6.5. Hibridaciones *in situ* en embriones enteros en estadio E6.5 muestran la falta de expresión de los genes específicos de células madre de trofoblasto (TS) *Cdx2* (A, A'), *Esrrb* (B, B') y *Eomes* (C, C'), que se encuentran en embriones silvestres expresados en la región de trofoblasto adyacente al epiblasto. La expresión de genes propios de linajes diferenciados es normal. *Ascl2* (D, D') se expresa de forma normal en el cono ectoplacentario de los embriones *Dicer*^{-/-} y *Rhox5* (E, E') presenta en todos los mutantes un patrón normal en ectodermo extraembrionario, cono ectoplacentario y endodermo visceral extraembrionario.

El análisis de expresión de *Esrrb*, *Eomes*, *Ascl2* y *Rhox5* en un estadio posterior e inmediatamente previo a la muerte del embrión (E7.5) confirma para todos los embriones mutantes la falta específica de expresión de genes implicados en el mantenimiento de la multipotencia de las células TS (Fig. 3A-B'; *Esrrb*, n=5/5; *Eomes*, n=3/3), mientras que no parece haber un defecto de expresión de genes característicos de linajes diferenciados (Fig. 3C-D'; *Ascl2*, n=3/3; *Rhox5*, n=5/5). El patrón de expresión de *Eomes* tanto en el trofoblasto

como en el mesodermo embrionario previo a la gastrulación nos permite comprobar que a los 7.5 días de desarrollo la pérdida de expresión de este gen es específica del trofoblasto. Y por otro lado indica que si bien el epiblasto es capaz de especificar el mesodermo embrionario, existe un retraso en el desarrollo ya que el patrón de expresión de *Eomes* en estadio E7.5 corresponde al de un embrión de estadio E6.5, en el que la línea primitiva comienza a formarse y aún no se ha extendido (Fig. 3B').

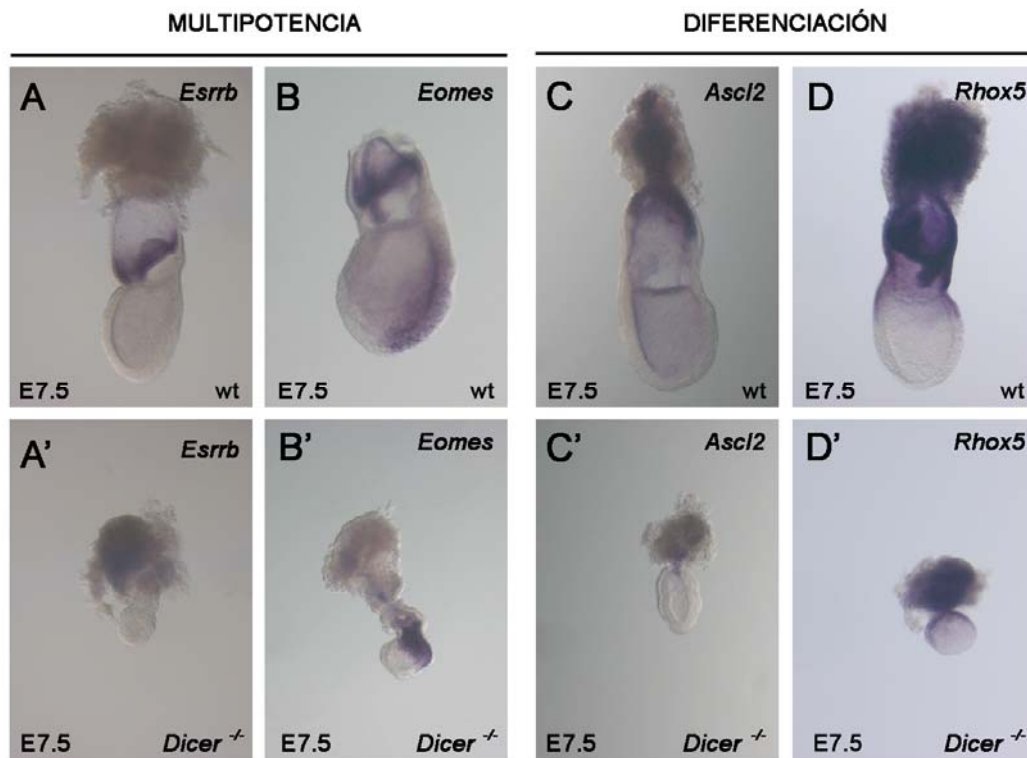


Figura 3. Los defectos específicos de células TS se mantienen en estadio E7.5. Hibridaciones *in situ* en embriones enteros en estadio E7.5 muestran de nuevo un defecto en la expresión de genes implicados en multipotencia (*Esrrb*, A, A' y *Eomes*, B, B') acompañado de una reducción general del ectodermo extraembrionario que se aprecia en todos los embriones mutantes. La falta de expresión de *Eomes* es específica de trofoblasto dado que se expresa en el mesodermo embrionario aunque con un patrón retrasado, correspondiente a un embrión en estadio E6.5. Los genes de diferenciación *Ascl2* (C, C') y *Rhox5* (D-D') se expresan de forma similar en los embriones mutantes y silvestres.

2.2 El fenotipo del trofoblasto es independiente del epiblasto

El mantenimiento de la población de células TS dentro del ectodermo extraembrionario depende en buena medida de la señalización por parte del epiblasto adyacente. Se han descrito dos vías principales que actúan en este sentido: la señalización por Nodal y la señalización por Fgf4 (Guzman-Ayala et al., 2004; Rossant and Cross, 2001; Tanaka et al., 1998). El factor de crecimiento Fgf4 se expresa desde muy temprano en el epiblasto embrionario y difunde al ectodermo extraembrionario. Allí se une a su receptor Fgfr2 desencadenando una cascada de señalización que promueve la expresión de *Cdx2*, *Esrrb* y *Eomes* en las células TS y al mismo tiempo reprime la expresión *Ascl2*, que queda restringido a la zona de diferenciación del cono ectoplacentario (Rossant and Cross, 2001; Tanaka et al.,

1998). Del mismo modo, *Nodal* se expresa en el epiblasto y es procesado en el ectodermo extraembrionario por las convertasas *Spc1* y *Spc4*. *Nodal* procesado contribuye de forma similar a como lo hace *Fgf4* al mantenimiento de las células TS en el trofoblasto, y al mismo tiempo contribuye en el epiblasto al establecimiento del patrón de este tejido (Guzman-Ayala et al., 2004; Tam and Loebel, 2007).

Dado que el epiblasto de los embriones *Dicer*^{-/-} es morfológicamente anormal (ver apartado 1), quisimos saber si la falta específica de expresión de genes propios de células TS podría deberse a un efecto secundario de falta de señalización por parte del epiblasto. El análisis por hibridación *in situ* en embriones enteros mostró una expresión normal de los genes implicados en esta comunicación. Todos los embriones mutantes analizados presentan niveles normales de expresión de *Fgf4* en el epiblasto tanto a estadio E5.5 (n=4/4, datos no mostrados) como a estadio E6.5 (n=2/2), así como de su receptor *Fgfr2* en el trofoblasto (n=3/3) (Fig. 4A-B'). Del mismo modo, *Nodal* se expresa en el epiblasto de todos los embriones mutantes analizados (Fig. 4C, C'; n=4/4) y los genes que codifican para las proteasas *Spc1* y *Spc4* se expresan en el trofoblasto de forma normal (Fig. 4 D-E'; n=3/3, y n=2/2, respectivamente).

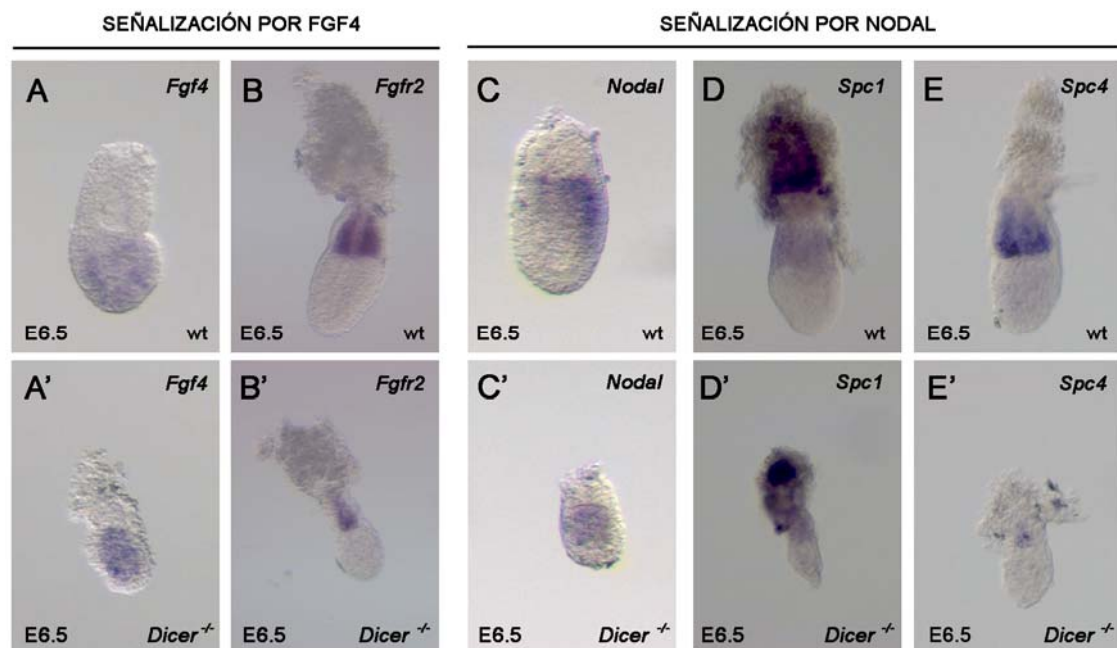


Figura 4. El defecto del trofoblasto no es debido a falta de señalización por parte del epiblasto. La hibridación *in situ* en embriones enteros muestra expresión normal de los genes implicados en la comunicación del trofoblasto y el epiblasto. Los genes implicados en la señalización por *Fgf4* se expresan de forma normal tanto en el epiblasto (*Fgf4*, A, A') como en el trofoblasto (*Fgfr2*, B, B') de embriones mutantes. El epiblasto de los embriones *Dicer*^{-/-} presenta expresión normal de *Nodal* (C, C'), mientras que los genes *Spc1* (D, D') y *Spc4* (E, E') están a su vez expresados de forma normal en el trofoblasto de embriones mutantes.

Estos resultados parecen indicar que el defecto en el mantenimiento de las células TS en los embriones mutantes para *Dicer* no es debido a defectos de expresión de genes implicados en la comunicación con el epiblasto.

2.3 El trofoblasto de los embriones *Dicer*^{-/-} presenta defectos de proliferación

Los embriones mutantes para *Dicer* presentan un defecto específico de pérdida de expresión de genes implicados en el mantenimiento de la población de células madre del trofoblasto. Como tales, las células TS han de tener una tasa de auto-renovación tal que les permita mantener un reservorio de células indiferenciadas. Ante la aparente pérdida o reducción de esta población celular existe la posibilidad de que esto sea debido a un aumento de la muerte celular o bien a una disminución en la proliferación. En cualquiera de estos dos supuestos el resultado sería una disminución en el número de células TS y por tanto una reducción tanto en la expresión de genes propios de este tipo celular dentro del ectodermo extraembrionario como en el tamaño de este propio tejido.

Estudiamos estas dos posibilidades en primer lugar analizando por ensayo de TUNEL la apoptosis en el trofoblasto de embriones *Dicer*^{-/-} tanto en estadio E5.5 como en E6.5. En ambos casos no se observó para ninguno de los embriones mutantes diferencias respecto a los silvestres en la tasa de apoptosis en el trofoblasto, tejido que ya de por sí en condiciones normales presenta una tasa de muerte muy baja (E5.5, n=6/6; E6.5, n=4/4; datos no mostrados).

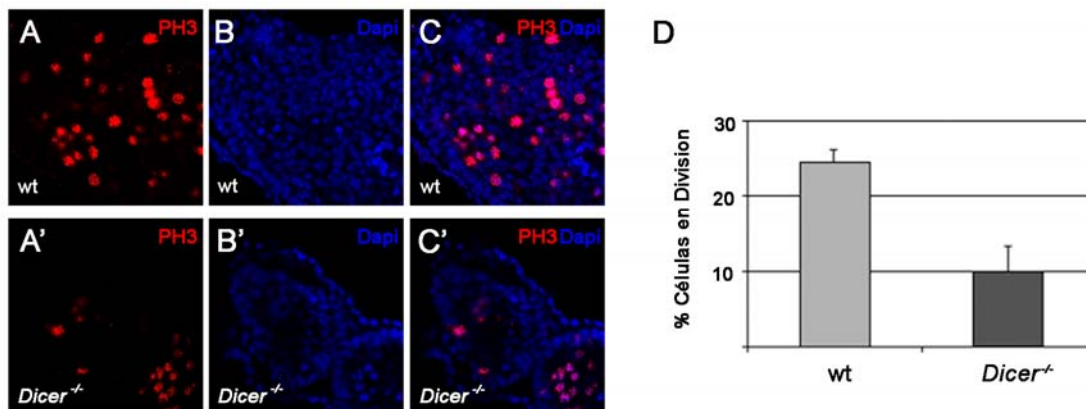


Figura 5. El trofoblasto de los embriones *Dicer*^{-/-} presenta un defecto de proliferación. La inmunodetección de PH3 muestra el número de células en división tras el tratamiento de 1h en nocodazol. La disminución en el número de células en división es aparente en el trofoblasto de los embriones mutantes respecto a los silvestres cuando se analizan en el microscopio confocal (A-C'). El recuento del número de células en división muestra una bajada significativa de la proliferación en el trofoblasto de embriones *Dicer*^{-/-} (D, contrastado con el estadístico T-student, p<0.05, las barras de error representan el Error de la Media).

Un posible defecto en la proliferación del trofoblasto se analizó mediante el recuento de las células en división en este tejido en estadio E6.5. Para ello hicimos un ensayo de inmunodetección de fosfo-histona 3 (PH3) en embriones previamente incubados durante una

hora en nocodazol, que acumula las células en metafase mitótica y permite así el recuento de un mayor número de células en división. Los resultados, mostrados en la figura 5, indican que en el ectodermo extraembrionario de los embriones *Dicer*^{-/-} existe una disminución estadísticamente significativa del porcentaje de células en división respecto al de los silvestres.

En conjunto, los resultados obtenidos del análisis *in vivo* de la falta de función de Dicer en el trofoblasto apuntan a una pérdida específica de expresión de genes implicados en el mantenimiento de la población de células TS a los 6.5 días de desarrollo, que culmina en una reducción del tejido del ectodermo extraembrionario apreciable en el estadio E7.5. Esta pérdida de células TS no parece deberse a un efecto secundario de falta de señalización por parte del epiblasto, ya que los genes encargados de esta comunicación se expresan de forma normal. Al contrario de lo que sucede con el mantenimiento de las células multipotentes, los procesos tempranos de diferenciación del trofoblasto no aparecen afectados por la ausencia de Dicer, siendo la expresión tanto de *Ascl2* como de *Rhox5* normales en todos los embriones mutantes analizados tanto a estadio E6.5 como a E7.5. Por último, la pérdida de células TS y posterior reducción del ectodermo extraembrionario viene acompañada por una bajada de proliferación en este tejido sin que la muerte celular se vea afectada.

ANÁLISIS *IN VITRO*: FENOTIPO DE LAS CÉLULAS TS-*Dicer*^{-/-}

Del análisis del fenotipo de embriones mutantes para *Dicer* se desprende un defecto en el mantenimiento de las células TS dentro del trofoblasto, así como un defecto de proliferación de este tejido. Dadas las dificultades asociadas al estudio *in vivo* decidimos utilizar un modelo *in vitro* para profundizar en el defecto de las células TS. Éstas se pueden derivar a partir de blastocistos y mantener en cultivo de manera indefinida, de modo que conservan su capacidad de auto-renovación y la multipotencia para poder diferenciarse en cualquiera de los linajes del trofoblasto y la placenta (Tanaka et al., 1998). De este modo, representan un modelo ideal para el estudio de defectos asociados al trofoblasto.

Al contrario de lo que sucede con las células madre embrionarias o ES, no se ha estudiado la falta de función de Dicer en ninguna línea de células TS. En este apartado describimos la caracterización fenotípica de las células TS mutantes para Dicer, generadas por derivación a partir de blastocistos *Dicer*^{fx/fx} y posterior delección del alelo *in vitro* (ver Materiales y Métodos).

3 Las células TS-*Dicer*^{-/-} tienen defectos en el mantenimiento de la multipotencia y la proliferación

Con el objetivo de analizar la falta de función de Dicer en células madre de trofoblasto derivamos dos líneas de células TS (TSL1-*Dicer*^{fx/fx} y TSL4-*Dicer*^{fx/fx}) a partir de blastocistos homocigotos para el alelo condicional *Dicer*^{fx/fx}. Ambas líneas celulares se comportan de forma

idéntica a líneas de células TS previamente descritas: se pueden mantener en cultivo de manera indefinida y conservan la capacidad de diferenciación (ver Materiales y Métodos) (Mak et al., 2002; Tanaka et al., 1998). Las células TS-*Dicer*^{fx/fx} correspondientes a ambas líneas se infectaron con el adenovirus AdCre-GFP para deletar *Dicer in vitro* y obtener así las células TSL1-*Dicer*^{-/-} y TSL4-*Dicer*^{-/-} respectivamente. Tras comprobar que ambas líneas se comportan de forma idéntica decidimos llevar a cabo todos los experimentos con la línea TSL4-*Dicer*^{fx/fx}, en lo sucesivo nos referiremos a esta línea como TS-*Dicer*^{fx/fx} o TS-*Dicer*^{-/-} tras la infección con AdCre-GFP. El sistema utilizado para obtener las células TS mutantes nos permite llevar a cabo un seguimiento del fenotipo inmediatamente después de la delección de *Dicer* y por tanto de la pérdida de los miRNAs maduros.

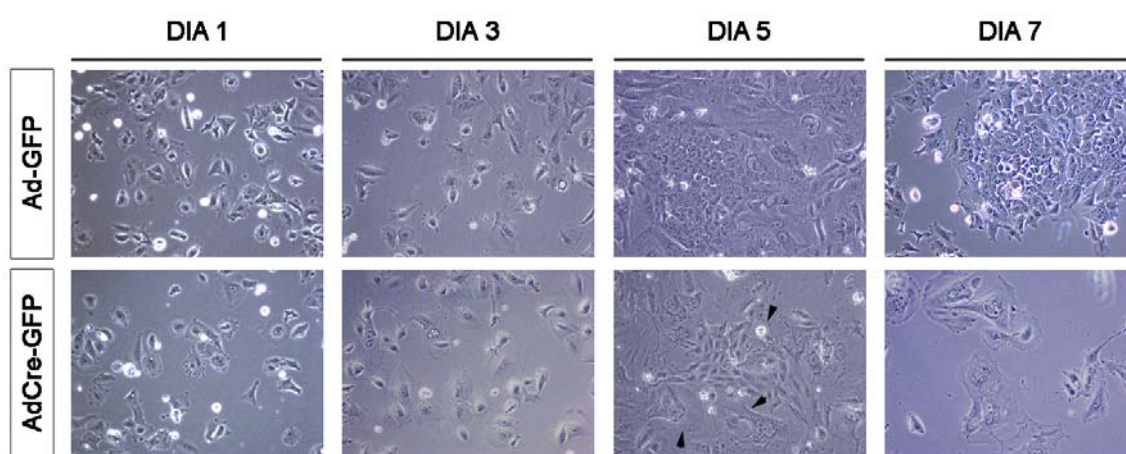


Figura 6. Las células TS-*Dicer*^{fx/fx} se diferencian en un periodo de 7 días tras la infección con AdCre-GFP. La infección de células TS-*Dicer*^{fx/fx} con el adenovirus AdCre-GFP que expresa la recombinasa Cre provoca la delección de *Dicer* en estas células. Transcurridas 24h (día 1) desde la infección las células infectadas con AdCre-GFP presentan igual morfología y que las infectadas con el adenovirus control AdGFP. Transcurridos tres días (día 3) siguen sin apreciarse diferencias entre células infectadas con ambos tipos de virus. A día 5 las células infectadas con AdCre-GFP han crecido menos, forman colonias más pequeñas y presentan un mayor porcentaje de diferenciación en células gigantes respecto a las infectadas con el virus control (las puntas de flecha señalan células gigantes rodeando una colonia de células TS). A día 7 tras la infección el número de células en el cultivo infectado con AdCre-GFP es mucho menor y todas las células TS se han diferenciado a células gigantes, mientras que las infectadas con AdGFP han crecido y formado colonias compactas de células TS indiferenciadas.

Las células TS-*Dicer*^{fx/fx} se infectaron con el adenovirus AdCre-GFP para deletar *Dicer* o con AdGFP como control negativo. A las 24 horas de la infección se seleccionaron en ambos casos las células positivas para GFP mediante citometría de flujo y éstas se mantuvieron en cultivo durante un total de seis días. La figura 6 refleja los cambios morfológicos observados en las células infectadas con AdCre-GFP respecto a las infectadas con el virus control AdGFP. Transcurridas 24 horas desde la infección y tras sembrar las células GFP positivas (día 1) no se aprecia ninguna diferencia notable en la morfología de las células infectadas con los dos tipos de virus. A los tres días de la infección sigue sin haber diferencias notables en morfología de las células, formación de colonias o número de células por placa. Sin embargo, a los 5 días de la infección comienzan a apreciarse en las células infectadas con AdCRE-GFP claros signos

de diferenciación alrededor de las colonias de TS. Así, es notable el incremento en el número de células de gran tamaño y con grandes núcleos semejantes a un tipo celular diferenciado que son las células gigantes de trofoblasto. Además, las colonias aparecen menos compactas, más pequeñas y con células TS ligeramente más grandes comparadas con las colonias observadas en las células infectadas con AdGFP. En este punto se aprecia además un menor número de células en las placas infectadas con AdCre-GFP. Transcurridos siete días las células TS han desaparecido del cultivo infectado con AdCre-GFP y en su lugar sólo quedan células gigantes de gran tamaño y en estado muy avanzado de diferenciación. Así pues, la delección de *Dicer* en células TS lleva de forma irrevocable a la diferenciación en un periodo de siete días, observándose los primeros indicios del fenotipo entre los días 3 y 5 después de la infección con AdCre-GFP.

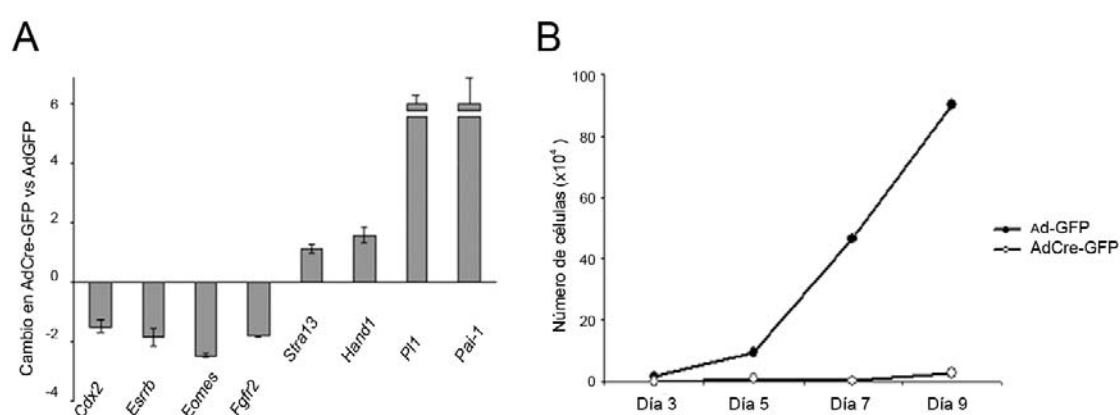


Figura 7. Las células TS-*Dicer*^{fx/fx} tienen defectos de mantenimiento de la multipotencia y auto-renovación. (A) El análisis de expresión por PCR cuantitativa a los siete días de la infección refleja una disminución en la expresión de marcadores de pluripotencia (*Cdx2*, *Esrrb*, *Eomes*, *Fgfr2*) y un aumento en expresión de marcadores de diferenciación (*Stra13*, *Hand1*, *Pl1*, *Pai-1*) en las células infectadas con AdCre-GFP respecto a las infectadas con AdGFP (el cambio se refiere en cada caso al número de veces que aumenta o disminuye la expresión). (B) La curva de crecimiento realizada sobre células TS infectadas con AdCre-GFP y AdGFP refleja un claro defecto de proliferación de las células mutantes frente a las silvestres.

El fenotipo de diferenciación observado morfológicamente se confirmó mediante la realización de PCR cuantitativa a los 7 días de la infección con AdCre-GFP (Fig. 7A). La bajada en expresión de los genes fundamentales en el mantenimiento de la auto-renovación y multipotencia de las células TS -como son *Cdx2*, *Esrrb*, *Eomes* y *Fgfr2*- implica una clara pérdida de este tipo celular (Rielland et al., 2008). Asimismo, la diferenciación de las células TS se confirma por el aumento de expresión de genes propios de linajes diferenciados tales como *Stra13*, *Hand1*, y en mayor medida *Pl1* y *Pai-1* (Erlebacher et al., 2004; Harvey et al., 1995; Hughes et al., 2004). En este caso es muy clara la subida de la prolactina *Pl1*, producida por las células gigantes de trofoblasto. La amplificación por PCR cuantitativa a los 4 días de la infección mostró la ausencia de mRNA de *Dicer* en células infectadas con AdCre-GFP

(disminución de 300 veces en comparación con células infectadas con Ad-GFP, datos no mostrados).

Las curvas de crecimiento llevadas a cabo a lo largo de los siete días de la infección muestran un claro defecto en la proliferación de las células infectadas con AdCre-GFP de forma similar a lo observado a simple vista en los cultivos (Fig. 7B).

El fenotipo de falta de función de Dicer observado para las células TS implica una bajada drástica de la proliferación acompañada de diferenciación, de modo que en un periodo de seis a siete días tras la delección de *Dicer* las células TS mutantes han desaparecido del cultivo y en su lugar quedan únicamente células en estado avanzado de diferenciación.

4 Cambios en RNA mensajero en células TS tras la delección de *Dicer*

Una vez observado a grandes rasgos el efecto de falta de función de Dicer en células TS, decidimos ahondar más en el fenotipo mediante la realización de un array de expresión de RNA mensajero en las células TS-*Dicer*^{-/-}. Dada la naturaleza del fenotipo observado resulta imposible generar una línea estable de células mutantes, por lo el análisis se llevó a cabo en células recogidas a días 4 y 6 tras la infección con AdCre-GFP. De este modo podemos analizar los cambios en expresión génica en el momento en que comenzamos a ver el fenotipo (día 4) y en el momento en que el fenotipo ya está establecido (día 6).

A los cuatro días de la infección hay un total de 1545 genes cuya expresión está alterada en las células infectadas con AdCre-GFP respecto a las infectadas con AdGFP. De ellos un 54% tienen un aumento en expresión (Fig. 8A). A día 6 tras la infección hay aproximadamente el doble de genes alterados en expresión (3115), que se ve aumentada en un 56% de los casos (Fig. 8B). La pérdida de función de Dicer supone un impedimento en la maduración de los miRNAs (Babiarz et al., 2008). Puesto que nuestro interés se centra en el estudio de la función de estos reguladores en el trofoblasto y las células TS, sólo consideraremos en análisis posteriores aquellos genes cuya expresión se incrementa tras la retirada de Dicer, ya que serán éstos los principales candidatos a ser reprimidos por miRNAs. Atendiendo únicamente a los genes cuya expresión aumenta tras la delección de *Dicer*, vemos que la mayor parte de los que varían a día 4 (88%) lo hacen también a día 6 (Fig. 8C). Así pues, decidimos centrar el análisis global de la variación en RNA mensajero en aquellos genes cuya expresión se encuentra aumentada a los cuatro días de la infección. Dado que es en este momento cuando comienza a apreciarse el fenotipo es de esperar que éstos sean los primeros genes que se vean alterados y que representen aquellos directamente regulados por miRNAs en células TS.

Como una primera aproximación al análisis de los genes con expresión aumentada a día 4 tras la infección llevamos a cabo un análisis de "Gene Ontology" (GO) sobre aquellos genes cuya variación en expresión es superior a un incremento de 1.3 veces con un p-valor inferior a 0,05 (Al-Shahrour et al., 2006; Al-Shahrour et al., 2005). En la figura 8D se representan los términos

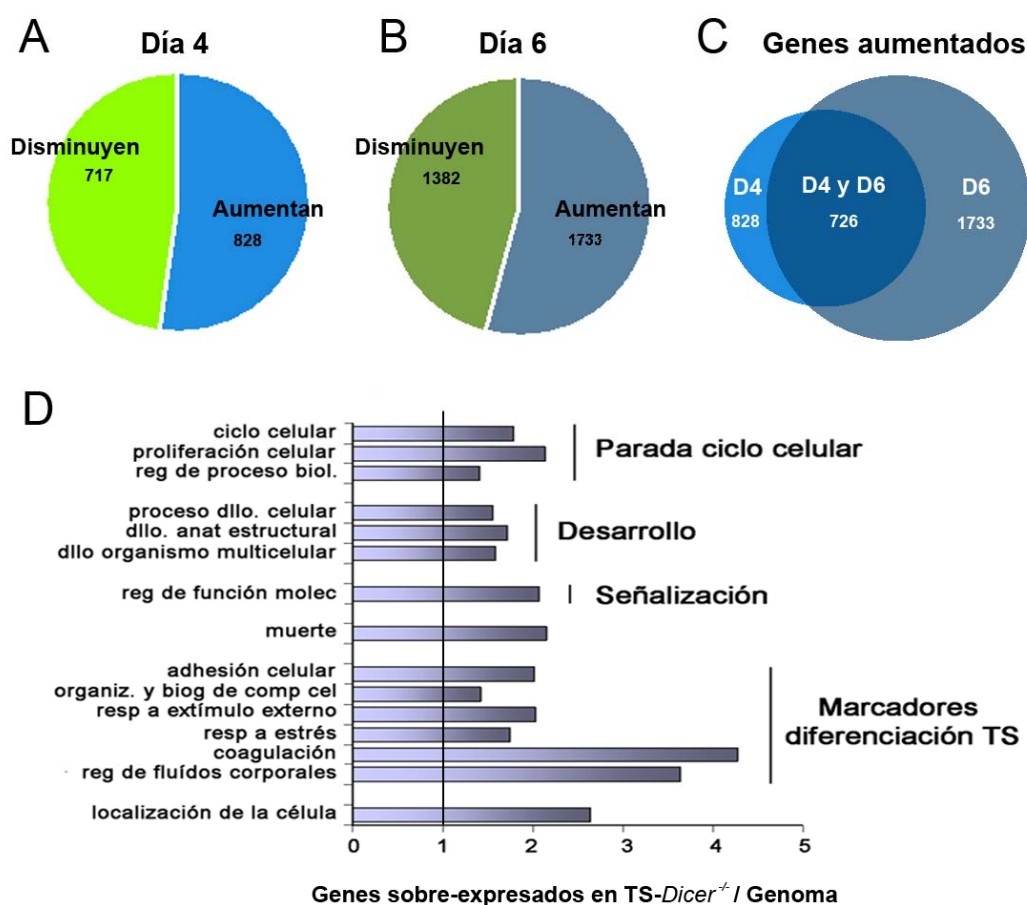


Figura 8. Variación en expresión génica en las células TS-Dicer^{-/-}. Los gráficos representan la variación en expresión génica en células infectadas con AdCre-GFP respecto a las infectadas con AdGFP ($p < 0.05$, Cambio > 1.3 veces). **(A)** A los cuatro días de la infección con AdCre-GFP se aprecia una alteración en expresión de 1545 genes en las células TS-Dicer^{-/-}, con un 54% de genes que incrementan su expresión. **(B)** A los seis días tras la infección con AdCre-GFP un total de 3115 genes varían en su expresión, de los cuales un 56% aumentan. **(C)** La gran mayoría (88%) de los genes sobre-expresados a los cuatro días de la infección con AdCre-GFP (D4) lo están también a los seis días (D6). **(D)** Análisis de “Gene Ontology” (GO) al nivel superior de proceso biológico (p corregido < 0.05), realizado sobre los genes sobre-expresados a día cuatro de la infección con AdCre-GFP ($p < 0.05$, Cambio > 1.3 veces). El resultado muestra un total de 15 términos sobre-representados en el grupo de genes sobre-expresados en las células mutantes respecto a la proporción en que estos términos se encuentran en el genoma. Estos términos se agrupan en 5 grandes categorías en función de los términos de nivel inferior (más específicos) que representan.

de GO correspondientes a un nivel 3 de Proceso Biológico que se encuentran sobre-representados en este set de genes con respecto al genoma (p -valor corregido $< 0,05$). Si analizamos en detalle los genes a los que se refiere cada uno de los 15 términos que aparecen como sobre-representados en el nivel más alto de GO (nivel 3), podemos clasificarlos 6 grandes grupos o procesos: parada de ciclo celular, desarrollo, señalización, muerte, diferenciación de células TS y otros. Estos resultados apuntan a un fenotipo similar al observado hasta el momento, en el que existe una diferenciación de células TS acompañada de parada de crecimiento. A su vez indica la alteración en genes implicados tanto en vías de señalización como en desarrollo, lo cual podría ser en parte la causa del fenotipo observado.

Tabla 1. Variación de genes de multipotencia y diferenciación tras la delección de *Dicer* en TS

Gen	Día 4		Día 6		Linaje	Ref.
	Cambio	Reg. en TS- <i>Dicer</i> ^{-/-}	Cambio	Reg. en TS- <i>Dicer</i> ^{-/-}		
<i>Esrrb</i>	1.6	Disminuye	1.6	Disminuye	TS	(Rielland et al., 2008)
<i>Eomes</i>	1.5	Disminuye	1.3	Disminuye	TS	(Rielland et al., 2008)
<i>Cdx2</i>	1.4	Disminuye	1.2	Disminuye	TS	(Rielland et al., 2008)
<i>Sox2</i>	1.7	Disminuye	1.8	Disminuye	TS	(Rielland et al., 2008)
<i>Spc4</i>	1.6	Disminuye	1.5	Disminuye	TS	(Guzman-Ayala et al., 2004)
<i>Id2</i>	1.3	Disminuye	1.4	Disminuye	TS	(Hemberger and Cross, 2001)
<i>Elf5</i>	1.9	Disminuye	2.1	Disminuye	TS	(Ng et al., 2008)
<i>Rhox6</i>	2.3	Aumenta	3.5	Aumenta	ExE, EPC	Nuestros resultados, no publicado
<i>Rhox9</i>	2.2	Aumenta	2.6	Aumenta	ExE,EPC	Nuestros resultados, no publicado
<i>Vgll3</i>	1.9	Aumenta	3.1	Aumenta	EPC	Nuestros resultados, no publicado
<i>Plac1</i>	2.2	Aumenta	3.3	Aumenta	TGCs	(Hughes et al., 2004)
<i>Mmp9</i>	2.2	Aumenta	2.4	Aumenta	TGCs	(Harvey et al., 1995)
<i>Gata2</i>	1.4	Aumenta	1.8	Aumenta	TGCs	(Ma et al., 1997)
<i>Met</i>	1.6	Aumenta	2.2	Aumenta	TGCs	(Hemberger and Cross, 2001)
<i>Vegfa</i>	1.7	Aumenta	1.8	Aumenta	TGCs	(Cross, 2000)
<i>Vegfc</i>	1.3	Aumenta	1.5	Aumenta	TGCs	(Cross, 2000)
<i>Adm</i>	1.4	Aumenta	1.6	Aumenta	TGCs	(Li et al., 2006)
<i>Plau</i>	2.5	Aumenta	4.2	Aumenta	TGCs	(Harvey et al., 1995)
<i>Pl1</i>	2.9	Aumenta	4.0	Aumenta	TGCs	(Hughes et al., 2004)
<i>Timp3</i>	2.1	Aumenta	2.8	Aumenta	TGCs	(Harvey et al., 1995)
<i>Itgav</i>	1.4	Aumenta	1.4	Aumenta	Laberinto	(Hemberger and Cross, 2001)
<i>Egfr</i>	1.3	Aumenta	1.0	-	Laberinto	(Hemberger and Cross, 2001)
<i>Vhlh</i>	1.3	Aumenta	1.8	Aumenta	Laberinto	(Hemberger and Cross, 2001)
<i>Tead3</i>	1.3	Aumenta	1.4	Aumenta	Laberinto	(Jacquemin et al., 1998)
<i>Pdgfrb</i>	1.9	Aumenta	2.3	Aumenta	Laberinto	(Hemberger and Cross, 2001)
<i>Pdgfb</i>	1.5	Aumenta	1.6	Aumenta	Laberinto	(Hemberger and Cross, 2001)
<i>Peg10</i>	2.1	Aumenta	2.4	Aumenta	Laberinto, TGCs	(Hemberger, 2007)
<i>Junb</i>	1.4	Aumenta	1.7	Aumenta	TGCs, Corion	(Hemberger and Cross, 2001)
<i>Snai1</i>	1.4	Aumenta	1.8	Aumenta	Espongiotrofoblasto	(Nakayama et al., 1998)
<i>Pparg</i>	1.3	Aumenta	1.6	Aumenta	Fusión corio-alant	(Hemberger and Cross, 2001)
<i>Cdh3</i>	1.3	Aumenta	1.3	Aumenta	ExE, EPC, TGCs	(Nose and Takeichi, 1986)

Índice: ExE-ectodermo extraembrionario; EPC-cono ectoplacentario;TGC-células gigantes de trofoblasto; Fusión Corio-alant- fusión corio-alantoidea

Teniendo en cuenta las agrupaciones resultantes del análisis de GO decidimos centrarnos en el estudio más en detalle de los genes implicados en estos procesos.

En primer lugar analizamos aquellos genes considerados como marcadores del fenotipo de pérdida de multipotencia y diferenciación de las células TS. En la tabla 1 se muestra cómo en las células TS-*Dicer*^{-/-} existe una bajada en expresión de una serie de genes propios de células indiferenciadas (parte superior, en rojo) y subida en expresión de genes propios de células diferenciadas pertenecientes a los distintos linajes del trofoblasto y la placenta (parte inferior, en azul). Esto confirma una vez más lo ya observado previamente en cuanto al fenotipo de pérdida de multipotencia y diferenciación de las células TS en ausencia de *Dicer*.

4.1 Rutas de señalización implicadas en la diferenciación

Dado que el análisis de GO indica una alteración en ciertas vías de señalización decidimos focalizar nuestra atención sobre aquellas rutas que se encuentran directamente implicadas en los procesos de mantenimiento de la multipotencia y auto-renovación de las células TS. En este caso ampliamos la búsqueda a genes que pudieran resultar interesantes por participar en éstas vías aunque su variación fuera inferior a 1.3 veces, pero siempre estadísticamente significativa.

4.1.1 Señalización por Fgf y TGFbeta

La señalización por Fgf4 a través de la fosforilación de Erk1/2 es absolutamente necesaria para el mantenimiento de las células TS indiferenciadas, por lo que cualquier inhibición de esta vía resultará en un fenotipo de diferenciación (Tanaka et al., 1998). En las células TS-*Dicer*^{-/-} hay dos genes que codifican para sendos inhibidores de la vía de Fgf (*Rasa2* y *Dusp9*) cuya expresión está aumentada a los cuatro días de la infección (Tabla 2) (Christie et al., 2005; Patterson et al., 2009; Yarwood et al., 2006). Del mismo modo, la expresión del gen que codifica para el receptor de Fgf4 en células TS (*Fgfr2*) (Rossant and Cross, 2001) se encuentra ligeramente disminuida a los cuatro días. En conjunto, la alteración en expresión de estos tres genes podría implicar una alteración en la vía de señalización por Fgf4.

Tabla 2. Alteración de genes implicados en señalización por Fgf4

Efecto sobre señalización por FGF	Gen	Día 4 post-infección		Día 6 post-infección	
		Cambio	Regulación en TS- <i>Dicer</i> ^{-/-}	Cambio	Regulación en TS- <i>Dicer</i> ^{-/-}
Inhibe	<i>Rasa2</i>	1.4	Aumenta	2.0	Aumenta
Inhibe	<i>Dusp9</i>	1.5	Aumenta	1.6	Aumenta
Potencia	<i>Fgfr2</i>	1.2	Disminuye	1.1	Disminuye

La señalización a través de la activación de Smad2/4 mediada por factores de la familia TGFbeta actúa de manera similar a como lo hace Fgf4, manteniendo las células TS en un estado indiferenciado y promoviendo la proliferación {Erlebacher, 2004 #25; Guzman-Ayala, 2004 #24}. En este sentido, no hemos encontrado ningún inhibidor de esta ruta cuya expresión se encuentre aumentada en las células mutantes de modo que pudiera ser causante del

fenotipo. Sin embargo, se ha descrito un efecto por el cual un exceso de señalización por factores de la familia TGFbeta en condiciones de baja señalización por Fgf4 potencia una parada en la proliferación de las células TS. La variación en expresión génica asociada a este proceso en concreto fue descrita por Erlebacher y colaboradores en 2004 {Erlebacher, 2004 #25} y se encuentra indicada en la tabla 3 (columna “Regulación por Aumento de TGFbeta”), en la que se indica a su vez la variación en expresión génica observada para las células TS-*Dicer*^{-/-}. Aparte de las variaciones indicadas en la tabla 3 es de destacar el aumento en expresión observado en el gen que codifica para el receptor de Activina (miembro de la familia de factores TGFbeta) *Acvrl1* (*Alk1*), que tiene un cambio de 1.3 veces a los cuatro días de la infección.

Tabla 3. Alteración de genes implicados en diferenciación por aumento de señalización por TGFbeta.

Reg. por Aumento de TGFbeta*	Gen	Día 4 post-infección		Día 6 post-infección	
		Cambio	Regulación en TS- <i>Dicer</i> ^{-/-}	Cambio	Regulación en TS- <i>Dicer</i> ^{-/-}
Aumenta	<i>Serpine1</i>	2.1	Aumenta	3.3	Aumenta
Aumenta	<i>Mmp9</i>	2.2	Aumenta	2.4	Aumenta
Aumenta	<i>Smad7</i>	1.1	-	1.3	Aumenta
Disminuye	<i>Myc</i>	1.4	Disminuye	1.4	Disminuye
Disminuye	<i>Id2</i>	1.3	Disminuye	1.4	Disminuye

* Regulación basada en el trabajo de Erlebacher y colaboradores (2004)

4.1.2 Otros factores implicados en diferenciación

La unión de la citoquina Lif (del inglés, “Leukemia Inhibitory Efactor”) a su receptor activa una vía de señalización que tiene como resultado la activación tanto de Stat3 como de Jak1. La señalización a través de esta ruta se ha visto implicada en la diferenciación hacia células gigantes de trofoblasto tanto en células TS como en la línea de coriocarcinoma Rcho-1 (Takahashi et al., 2003; Takahashi et al., 2008). Los datos obtenidos del array de mRNA en células TS-*Dicer*^{-/-} indican un aumento en la expresión tanto de *Lif* como del gen que codifica para su co-receptor *Il6st* (cambio de 1.5 y 2.0 veces respectivamente). Esto podría suponer un aumento en la señalización por esta vía que en último término potenciaría la diferenciación de las células TS hacia células gigantes.

Entre los genes cuya expresión aumenta en ausencia de *Dicer* hemos encontrado otros dos interesantes: *Col4a1* y *Col4a2*, que codifican para dos subtipos de ColágenoIV. Esta es una proteína de matriz extracelular cuyo aumento en expresión es capaz de provocar cierta reprogramación de células ES hacia el linaje de TS. En este caso, resultados obtenidos por Schenke-Layland y colaboradores (2007) indican que este proceso va acompañado de la sobre-expresión de genes propios de linajes diferenciados de trofoblasto. Tanto *Col4a1* como *Col4a2* presentan un incremento en expresión (de 1.4 y 1.3 veces respectivamente) en las células TS tras la retirada de *Dicer*. Además, cuatro de los genes descritos como sobre-expresados en respuesta a este aumento – *Rhox6*, *Rhox9*, *Plac1* y *Cdh3* (Schenke-Layland et

al., 2007)- varían también en este sentido (ver tabla 1). Así pues, un aumento en Colágeno IV podría estar colaborando en la expresión de genes propios de linajes diferenciados de trofoblasto.

4.1.3 Regulación por Cdkn1c

Las células gigantes de trofoblasto se caracterizan por ser poliploides debido a un proceso de endo-reduplicación que tiene lugar cuando comienza el programa de diferenciación hacia este tipo celular. Este proceso está controlado fundamentalmente por los inhibidores de proteína quinasa dependiente de ciclina Cdkn1c (p57) y Cdkn1a (p21). Recientemente se ha descrito en detalle cómo el aumento en expresión de *Cdkn1c* debido a la falta de señalización por Fgf4 se encuentra íntimamente ligado a la diferenciación de las células TS hacia células gigantes (Ullah et al., 2008). La expresión de *Cdkn1c* se encuentra significativamente elevada en las células TS-*Dicer*^{-/-} a los 4 y 6 días tras la infección (tabla 4). Si atendemos a los genes cuya variación en expresión se ha visto asociada a la diferenciación provocada por el aumento de Cdkn1c (tabla 4, columna Regulación en Diferenciación) vemos que la mayor parte de ellos varían en igual sentido en las células TS tras la delección de *Dicer*. Esto apunta a un incremento de esta vía que podría estar asociada al aumento de diferenciación en las células mutantes.

Tabla 4. Variación génica asociada al aumento de expresión de *Cdkn1c*

Reg. en Diferenciación*	Gen	Día 4 post-infección		Día 6 post-infección	
		Cambio	Regulación en TS- <i>Dicer</i> ^{-/-}	Cambio	Regulación en TS- <i>Dicer</i> ^{-/-}
Aumenta	<i>Cdkn1c</i>	1.5	Aumenta	2.0	Aumenta
Aumenta	<i>Cdkn1a</i>	1.9	Aumenta	2.3	Aumenta
No cambia	<i>Cdkn1b</i>	1.1	-	1.0	-
Disminuye	<i>Cdc2a</i>	1.3	Disminuye	1.6	Disminuye
Disminuye	<i>Chek1</i>	1.3	Disminuye	1.6	Disminuye
No cambia	<i>Gmnn</i>	1.3	Disminuye	1.3	Disminuye
No cambia	<i>Ccnb1</i>	1.3	Disminuye	1.9	Disminuye
Disminuye	<i>Ccna2</i>	1.3	Disminuye	1.6	Disminuye
Disminuye	<i>Ccne1</i>	1.2	Disminuye	1.1	-
Disminuye	<i>Atr</i>	1.2	Disminuye	1.4	Disminuye

* Regulación basada en el trabajo de Ullah y colaboradores (2008)

Por otro lado, es de destacar que existe a su vez un aumento en la expresión del gen codificante para la quinasa dependiente de ciclina Cdk6, con un cambio de 1.9 veces a los cuatro días tras la infección. Cdk6 está implicada a su vez en la represión de miembros de la familia del retinoblastoma (Malumbres and Barbacid, 2009), que sería necesaria en cierto modo para que las células puedan endo-reduplicar su DNA.

4.2 Ruta de señalización implicada en la bajada de proliferación

Dado que el fenotipo observado tras la delección de *Dicer* es claramente de bajada de proliferación, y atendiendo al análisis de GO previamente realizado, resulta obvia la búsqueda de genes implicados en salida de ciclo celular cuya expresión pueda estar aumentada en las

células TS-*Dicer*^{-/-}. En este sentido, son tres los factores que hemos encontrado asociados a la salida de ciclo celular cuya expresión está notablemente aumentada en las células TS-*Dicer*^{-/-}. Los tres se encuentran implicados en el paso de fase G1 a S del ciclo celular (Tabla 5). Estos son el supresor tumoral Lats2 (Li et al., 2003), el inhibidor de kinasa dependiente de ciclina Cdkn1a (p21) y el miembro de la familia de retinoblastoma Rbl2 (Blagosklonny and Pardee, 2002). Las proteínas de la familia del retinoblastoma colaboran en la represión de genes asociados con la síntesis de DNA, mientras que Lats2 y Cdkn1a potencian la función del retinoblastoma por medio de la inhibición de represores del éste.

Tabla 5. Variación en expresión de genes implicados en ciclo celular

Efecto sobre proliferación	Gen	Día 4 post-infección		Día 6 post-infección	
		Cambio	Regulación en TS- <i>Dicer</i> ^{-/-}	Cambio	Regulación en TS- <i>Dicer</i> ^{-/-}
Inhibe	<i>Cdkn1a</i>	1.9	Aumenta	2.3	Aumenta
Inhibe	<i>Lats2</i>	1.3	Aumenta	1.6	Aumenta
Inhibe	<i>Rbl2</i>	1.3	Aumenta	1.6	Aumenta

Los resultados obtenidos a partir del análisis de los datos del array de RNA mensajero en células TS mutantes apuntan en la misma dirección que los previamente obtenidos tanto *in vivo* como *in vitro*. Esto es, un fenotipo de diferenciación reflejado en el aumento de marcadores de linajes diferenciados de trofoblasto y disminución de aquellos propios de células TS. Además, los datos del array ponen de manifiesto la alteración en expresión de genes cuyos productos participan en vías de señalización importantes para el mantenimiento tanto de la multipotencia como de la proliferación de las células TS.

EXPRESIÓN DE miRNAs EN CÉLULAS TS Y EMBRIÓN TEMPRANO

Hasta el momento hemos descrito el fenotipo de falta de función de Dicer tanto en el trofoblasto de embriones en estadios previos a la gastrulación como en células TS. Sin embargo, el propósito inicial del trabajo es abordar los efectos de la falta de miRNAs funcionales, utilizando la delección de *Dicer* como herramienta para impedir su maduración. Una vez descrito el comportamiento tanto del trofoblasto como de las células TS en cultivo en ausencia de miRNAs funcionales se hace necesario conocer cuáles se expresan en estos linajes. Hasta el momento sólo se encuentra descrito su perfil de expresión en las células ES (Houbaviy et al., 2003), pero se desconoce el de células TS o el de embriones en estadios tempranos de post-implantación.

En este apartado describimos los perfiles de expresión de miRNAs obtenidos para células TS, ES y fibroblastos embrionarios (MEFs, del inglés, "Mouse Embryonic Fibroblasts") y para embriones en estadios E5.5, E6.5, E7.5 y E8.5.

5 Perfil de expresión de miRNAs en células TS y en embrión temprano

La obtención de los perfiles de expresión en las distintas muestras se llevó a cabo mediante un array por PCR cuantitativa de 312 miRNAs bien caracterizados en ratón (ver Materiales y Métodos). El análisis se realizó sobre dos líneas diferentes para cada tipo celular, incluyendo en el caso de las TS la línea TS-*Dicer^{flx/flx}* derivada por nosotros y la línea previamente caracterizada TS-B1 (Mak et al., 2002).

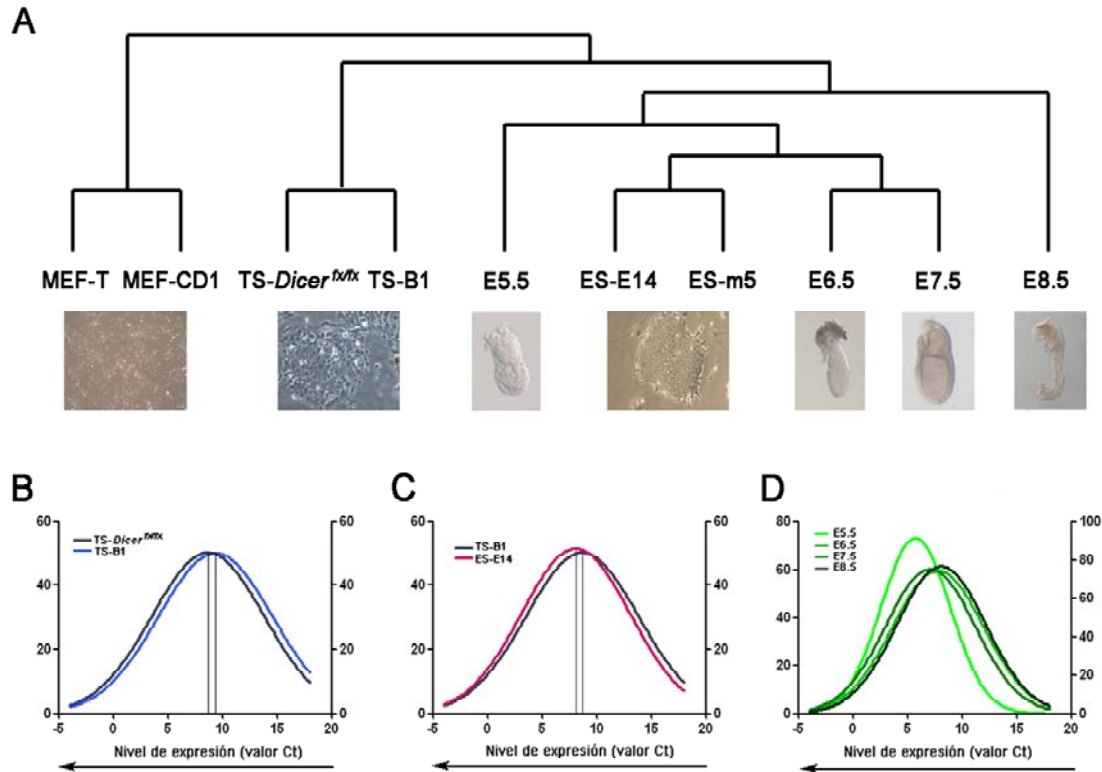


Figura 9. Comparación de los perfiles de expresión de miRNAs en células madre derivadas del blastocisto (ES y TS), MEFs y embriones tempranos. (A) Agrupamiento jerárquico de los perfiles de expresión de miRNAs en ES, TS, embriones y MEFs, utilizando los niveles de expresión (valores de Ct) de 312 miRNAs (coeficiente de correlación como medida de similitud y el ligamiento completo como método de agrupamiento). Las dos líneas de células ES presentan mayor similitud en su perfil de expresión de miRNAs con todas las muestras embrionarias, y quedan separados a su vez todos ellos de las líneas de células extraembrionarias TS. Las dos líneas de fibroblastos (MEF) agrupan juntas y forman un grupo aparte del resto de muestras. (B) La distribución de frecuencia realizada con los valores de expresión (valor de Ct) de las dos líneas de células TS muestra una distribución muy similar para ambas, con un nivel medio de expresión de Ct=9.0 en la línea TS-*Dicer^{flx/flx}* y Ct=8.7 en la línea TS-B1. La distribución de frecuencia para ambas líneas de células ES es igualmente muy similar (datos no mostrados). (C) Las células TS y ES presentan una distribución de niveles de expresión muy parecida, con un valor de expresión medio de Ct=8.4 en la línea ES-E14 y Ct=8.7 en la línea TS-B1. (D) Los embriones en estadio E5.5 presentan una distribución de frecuencia de niveles de expresión ligeramente desviada de la del resto de muestras embrionarias, con una media de expresión de Ct=6.4, superior tanto a la de los demás estadios como a la de las células ES y TS. Los embriones en estadio E6.5 a E8.5 muestran niveles de expresión muy similares entre sí con medias de expresión de Ct=8.0 (E6.5), Ct=7.5 (E=7.5) y Ct=8.4 (E=8.5).

Una vez obtenidos los perfiles de expresión para las células y los embriones llevamos a cabo un agrupamiento jerárquico de las muestras en función de la similitud en perfiles de expresión. El resultado de este agrupamiento se muestra en la figura 9A. Aquí vemos que las células ES tienen un perfil de expresión más similar al de los embriones tempranos, siendo los de estadio E8.5 los más alejados de las células madre embrionarias. Separadas de los embriones y células ES aparecen las células TS, y finalmente los fibroblastos embrionarios quedan como el grupo más alejado.

Las dos líneas utilizadas para cada tipo celular agrupan juntas, lo que confirma un perfil de expresión muy similar. Además, si analizamos las distribuciones de frecuencia en niveles de expresión para cada uno de los pares vemos éstas son también muy parecidas (Fig. 9B, datos no mostrados). En el caso de las células TS esto nos permite comprobar que la línea TS-*Dicer^{fx/fx}* que hemos derivado no tiene defectos en la expresión o procesamiento de miRNAs. Cuando comparamos los niveles de expresión de miRNAs entre TS y ES vemos que no hay gran diferencia entre ambos tipos celulares (Fig. 9C). Finalmente del mismo modo podemos ver que los embriones en estadio E5.5 tienen un nivel general de expresión de miRNAs superior tanto al del resto de embriones como al de las células ES y TS (con un valor medio de Ct de 6.37) (Fig. 9D).

Tabla 6. miRNAs altamente expresados en células TS

miRNA	Valor Ct TS-B1	Valor Ct TS- <i>Dicer^{fx/fx}</i>	Cambio TS-B1 vs ES-E14	Expresión predominante
mmu-miR-424	0.5	0.2	44.7	TS
mmu-miR-29a	2.9	1.4	9.1	TS
mmu-miR-532	3.7	3.8	9.9	TS
mmu-miR-365	3.9	5.3	8.9	TS
mmu-miR-322	4.4	3.6	35.7	TS
mmu-miR-467	0.9	2.8	8.5	TS, E5.5
mmu-miR-467b	0.9	2.3	8.4	TS, E5.5
mmu-miR-669a	1.0	2.3	4.0	TS, E5.5
mmu-miR-466	1.3	2.6	4.9	TS, E5.5
mmu-miR-30d	4.2	2.6	2.4	TS, E5.5
mmu-miR-22	4.4	5.3	3.1	TS, E5.5
mmu-let-7	4.9	4.3	12.1	TS, E5.5
mmu-let-7g	5.0	4.6	6.6	TS, E5.5
mmu-miR-26b	3.3	3.8	2.2	TS, ET
mmu-miR-301	2.8	3.1	1.8	TS, Embriones
mmu-miR-21	0.2	0.1	2.2	CM, E5.5
mmu-miR-7	3.1	5.3	-1.1	CM, E5.5
mmu-miR-295	-1.9	-1.4	-1.2	CM, ET
mmu-miR-294	-1.9	-1.6	-1.3	CM, ET
mmu-miR-292-3p	-1.3	-1.5	-1.8	CM, ET
mmu-miR-291-3p	0.2	0.8	1.2	CM, ET
mmu-miR-291b-5p	0.4	0.7	1.1	CM, ET
mmu-miR-92	0.8	0.9	-2.3	CM, ET
mmu-miR-292-5p	2.1	2.7	1.3	CM, ET
mmu-miR-18a	4.2	4.7	-1.4	CM, ET
mmu-miR-290	4.4	4.3	-1.8	CM, ET
mmu-miR-291-5p	4.5	4.7	-1.3	CM, E6.5, E7.5
mmu-miR-24	2.3	2.5	1.5	Células
mmu-miR-27a	2.2	2.6	1.5	Células
mmu-miR-93	2.6	2.2	-1.8	Células, ET
mmu-miR-30a-5p	2.9	1.9	1.6	Células, ET
mmu-miR-182	3.1	3.6	-1.2	Células, ET

Índice: CM-células madre (TS y ES); Embriones- E5.5 a E8.5; ET- E5.5 y E6.5; Células: TS, ES y MEF

La obtención de los perfiles de expresión de miRNAs en células TS nos ha permitido determinar cuáles son los que se encuentran más altamente expresados en este tipo celular y que por tanto pueden tener aquí una función. En la tabla 6 se encuentran listados todos aquellos miRNAs cuya expresión es elevada (determinado por un valor de Ct inferior a 5) y enriquecida en células TS respecto ES y/o embriones en los distintos estadios, excluyendo aquellos que se expresan de manera ubicua en todas las muestras analizadas. Se han incluido además en un último grupo aquellos cuya expresión es específica de células en cultivo, incluyendo en este caso ES, TS y MEF.

6 Perfil de expresión en relación a la organización genómica

Los miRNAs pueden encontrarse aislados en el genoma o formando parte de un grupo o cluster (Bartel, 2004). Al analizar su expresión vemos que en muchos casos los integrantes de un mismo cluster comparten un perfil similar en las distintas líneas celulares y estadios embrionarios. Así, en la figura 10 están representados cuatro ejemplos de clusters interesantes por su elevada expresión en células TS y/o en el embrión temprano. En ellos se aprecia un perfil parecido para todos los miRNAs del cluster analizados en nuestro array. El miRNA situado en posición 5' ha sido el utilizado para dar nombre al cluster en cada caso.

Los clusters miR-466 y miR-322/424 son interesantes por su expresión elevada y específica en células TS respecto a las ES (Fig. 10A, B). El cluster miR-290, previamente descrito como específico de ES y embrión en pre-implantación (Houbaviy et al., 2005), posee un perfil de expresión que lo sitúa como específico de células madre derivadas del blastocisto (TS y ES) así como de embrión temprano en pre y post-implantación, ya que su expresión decae en embriones en estadio E8.5 y no se expresa en fibroblastos embrionarios (Fig. 10C). Por su parte los miRNAs del cluster miR-106a se expresan preferentemente en las células ES y en el embrión en estadios tempranos, y no en células extraembrionarias (TS) o embrionarias tardías (MEFs) (Fig. 10D).

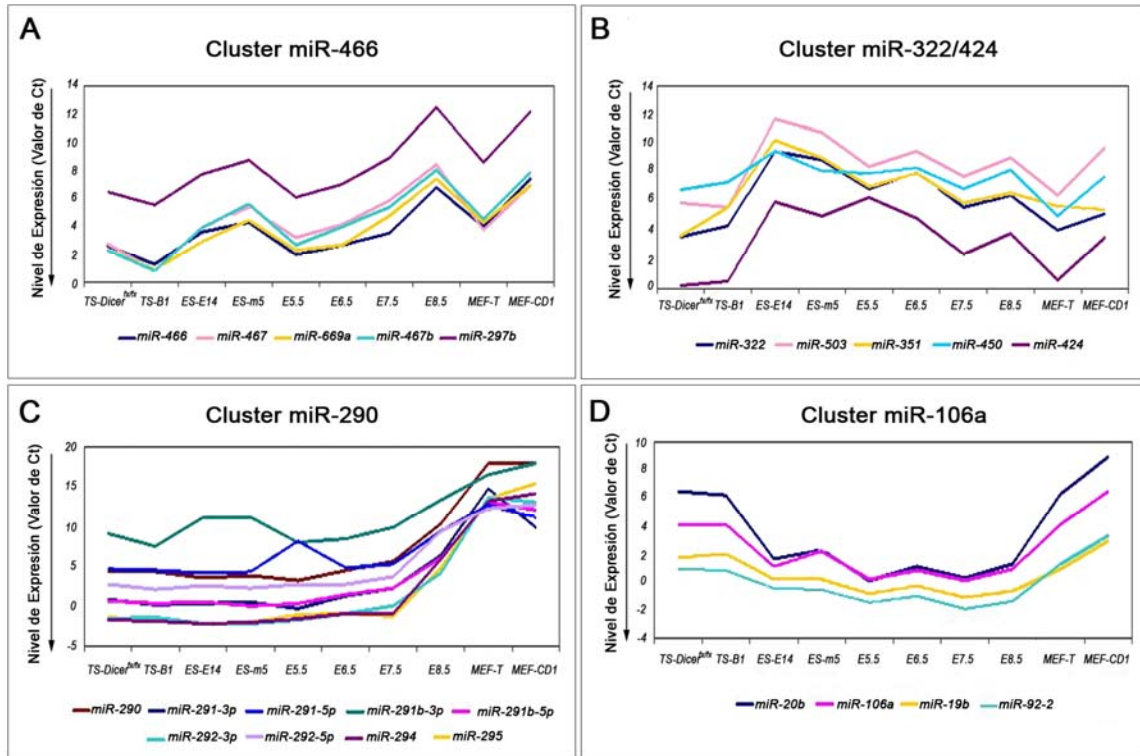


Figura 10. Perfiles de expresión de los clusters miR-466, miR-322/424, miR-290 y miR-106a en células TS, ES, embriones tempranos y MEFs. Nivel de expresión (valor de Ct) de los miRNAs integrantes de cada cluster analizados dentro del array de expresión de miRNAs en las distintas líneas celulares y muestras de embriones. Los miembros de un mismo cluster presentan perfiles de expresión en las distintas muestras muy similares entre sí. **(A)** Los miRNAs del cluster miR-466 presentan niveles de expresión más elevados (dados por un valor de Ct inferior) en las células TS y en los embriones tempranos (E5.5 y E6.5) que en células ES o embriones en estadios más avanzados. **(B)** el cluster miR-322/424 se puede considerar como específico de células TS con respecto a las muestras embrionarias y células ES, aunque los niveles de expresión son también elevados en una de las líneas de fibroblastos embrionarios (MEF-T). **(C)** El cluster miR-290 se encuentra altamente expresado en células madre derivadas del blastocisto (ES y TS) y en embrión temprano, disminuyendo su expresión a medida que el embrión diferencia un mayor número de tejidos. **(D)** El cluster miR-106a aparece aquí como específico de células ES y embriones hasta estadio E8.5, y tiene un nivel de expresión muy inferior tanto en células extraembrionarias (TS) como en embrionarias tardías (MEF).

6.1 Patrones de expresión de miRNAs restringidos

Los resultados previamente descritos reflejan el hecho de que determinados miRNAs se encuentran expresados de forma preferente en unos u otros tipos celulares y también asociados a un cierto grado de desarrollo o diferenciación del embrión temprano. En ciertos casos estos miRNAs poseen una organización genómica en cluster, compartiendo los miembros de un mismo cluster un perfil de expresión similar. De los cuatro ejemplos de clusters arriba mencionados dos de ellos merecen especial atención: el cluster miR-466 y el cluster miR-290, cuyo patrón de expresión detallamos a continuación.

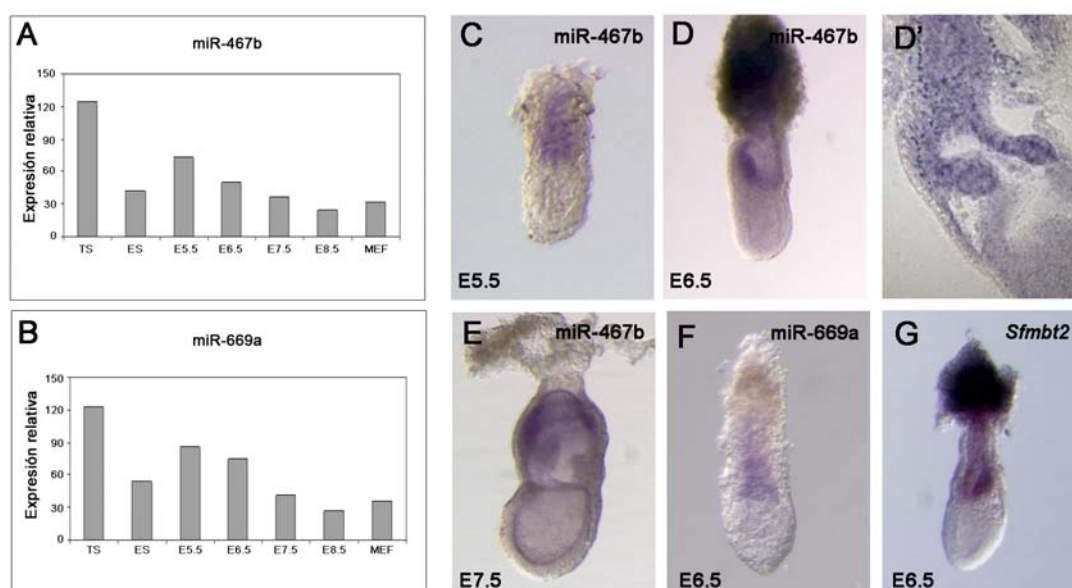


Figura 11. Patrón de expresión del cluster miR-466 en embrión temprano. (A-B) Los miRNAs miR-467b y miR-669a, pertenecientes al cluster 466, tienen niveles de expresión superiores en células extraembrionarias TS respecto al resto de muestras analizadas (representado en valores de expresión relativa obtenida por PCR cuantitativa, el valor en cada tipo celular corresponde a la media de expresión en ambas líneas). (C-D) La hibridación *in situ* en embriones enteros con la sonda LNA para miR-467b en estadios E5.5 y E6.5 muestra una restricción del patrón de expresión al linaje del trofoblasto, tanto en el ectodermo extraembrionario como en el cono ectoplacentario. No hay expresión en epiblasto ni en endodermo visceral embrionario o extraembrionario (D', vista al microscopio de sección del embrión en D). En estadio E7.5 miR-467b se expresa de forma específica en el corion, derivado del ectodermo extraembrionario (E). El otro miembro del cluster analizado, miR-669a, tiene un patrón de expresión idéntico al de miR-467b en estadio E6.5 (F). El cluster miR-466 se localiza íntegramente en el intrón 10 del gen *Sfm2t2*. El análisis por hibridación *in situ* de la expresión de *Sfm2t2* en embriones en E6.5 muestra un patrón de expresión igual al observado para los dos miembros del cluster miR-466 analizados (G).

El cluster miR-466 posee un perfil de expresión interesante a la hora de buscar miRNAs específicos no sólo de células TS sino también de trofoblasto, puesto que se encuentra altamente expresado tanto en células TS como en el embrión temprano. De acuerdo a las anotaciones de la base de datos miRBase (<http://microna.sanger.ac.uk/sequences/>) este cluster posee un total de 52 miRNAs, todos ellos localizados en el intrón 10 del gen *Sfm2t2* y en la misma orientación que el propio gen. Los miRNAs que integran este cluster pertenecen a diferentes familias, lo cual indica que reprimen sets distintos de mRNAs diana, sin embargo el perfil de expresión es similar en aquellos que hemos analizado. El análisis por hibridación *in situ* con sondas de LNAs para dos miRNAs pertenecientes a este cluster (miR-467b y miR-669a) muestra cómo efectivamente poseen un patrón de expresión restringido al trofoblasto en estadios tempranos (Fig. 11). El miRNA miR-467b se expresa en estadios E5.5 y E6.5 tanto en el ectodermo extraembrionario como en las zonas más diferenciadas del cono ectoplacentario, y a 7.5 días de desarrollo la expresión queda restringida al ectodermo extraembrionario del corion (Fig. 11C-E). El patrón observado para miR-669a en estadio E6.5 es idéntico al de miR-467b (Fig. 11F). Estos patrones coinciden con lo esperado según el perfil de expresión

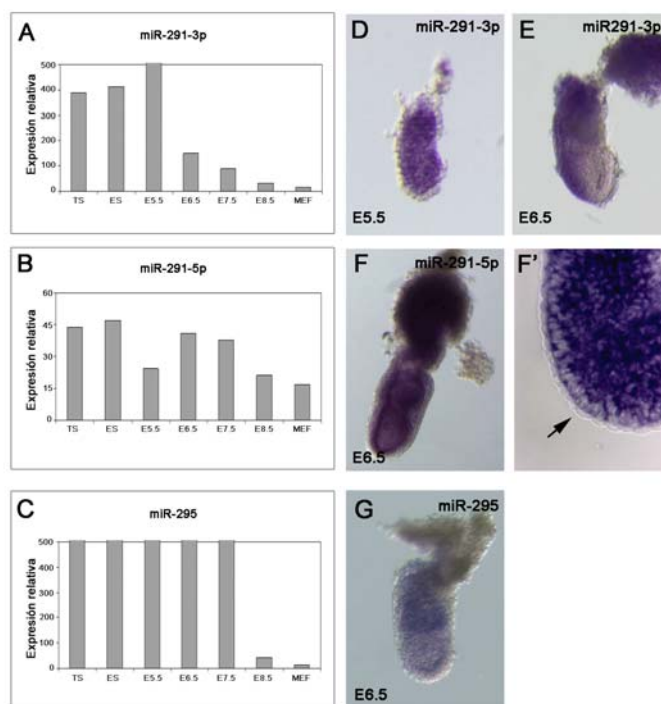
obtenido por PCR cuantitativa (Fig. 11A, B), y son además iguales al del gen *Sfmbt2* en cuyo intrón se localizan (Fig. 11G, (Frankenberg et al., 2007)).

El cluster miR-290 ha sido descrito como específico de ES y embrión en pre-implantación (Houbaviy et al., 2005) y nuestros resultados indican que se encuentra también altamente expresado en las células TS y en el embrión temprano en post-implantación, lo que podría ser indicativo de una función específica en los linajes que integran el embrión durante etapas muy tempranas del desarrollo. Por esta razón lo seleccionamos como el segundo cluster a analizar en detalle por hibridación *in situ* en embriones. Este cluster posee siete miRNAs y se encuentra localizado en el cromosoma 7 de ratón en una región intergénica. Al igual que sucede con el cluster miR-466, los miRNAs del cluster miR-290 no se encuentran todos dentro de la misma familia, siendo miR-291-3p, miR-294 y miR-295 los únicos del cluster que forman parte de la misma familia y que por tanto comparten los mismos mRNAs diana. Los perfiles de expresión obtenidos por PCR cuantitativa y los patrones de expresión observados por hibridación *in situ* con sondas LNAs para los miRNAs miR-291-3p, miR-291-5p y miR-295 se muestran en la figura 12. Los miRNAs miR-291-3p y miR-291-5p proceden los dos del mismo pre-miRNA pero de hebras complementarias. Los patrones de expresión observados para ambos son similares aunque no idénticos. miR-291-3p (Fig. 12D, E) se expresa en estadio E5.5 tanto en el epiblasto como en el trofoblasto, quedando únicamente excluido del endodermo visceral embrionario y extraembrionario. Sin embargo en estadio E6.5 la expresión persiste en ambos linajes pero parece encontrarse en el epiblasto restringida a un lado del embrión. En este caso resulta difícil determinar si se trata del lado anterior o posterior, y asimismo quedan dudas acerca de la expresión en el endodermo visceral de esta región del embrión. Es en cualquier caso interesante el hecho de que exista un patrón de expresión restringido a un lado del epiblasto en un momento en el que tiene lugar la especificación del eje antero-posterior. En estadio E7.5 no hemos detectado expresión de este miRNA (datos no mostrados). Al igual que miR-291-3p, el miRNA procesado a partir de la otra hebra, miR-291-5p, se expresa en estadio E6.5 en el trofoblasto y epiblasto. Sin embargo en este caso la expresión es fuerte y homogénea en ambos tejidos, siendo el endodermo visceral embrionario y extraembrionario el único tejido en el que no se aprecia expresión (Fig. 12F, F').

Por último hemos analizado el patrón de expresión en embriones en estadio E6.5 de miR-295, perteneciente a la misma familia que miR-291-3p. Al igual que los otros dos miRNAs previamente descritos, miR-295 se expresa en el embrión temprano tanto en el epiblasto como en el trofoblasto, pero no en el endodermo visceral. Sin embargo en este caso la expresión es uniforme dentro de cada tejido pero a su vez más fuerte en el linaje extraembrionario del trofoblasto. Así pues, pese a pertenecer a un mismo cluster y expresarse en idénticos tejidos, los miRNAs del cluster miR-290 parecen estar regulados de forma compleja de modo que difieren en los patrones de expresión dentro de cada linaje.

Figura 12. Patrón de expresión del cluster miR-290 en embrión temprano.

Los miRNAs miR-291-3p, miR-291-5p y miR-295, pertenecientes al cluster 290, tienen niveles de expresión superiores en células madre derivadas del blastocisto (TS y ES) y en embrión temprano respecto embriones en E8.5 y MEFs (representado en valores de expresión relativa obtenida por PCR cuantitativa, el valor en cada tipo celular corresponde a la media de expresión en ambas líneas) (A-C). La hibridación *in situ* en embriones enteros con sondas LNA muestra que miR-291-3p (D) se expresa en estadio E5.5 en todo el epiblasto y trofoblasto pero no en el endodermo visceral embrionario o extraembrionario; sin embargo su expresión queda restringida a un lado del epiblasto y trofoblasto en estadio E6.5 (E). No se detecta expresión de miR-291-3p por hibridación *in situ* en embriones en estadio E7.5 (datos no mostrados). miR-291-5p (F), sintetizado a partir de la hebra complementaria a miR-291-3p, se expresa en estadio E6.5 muy fuerte en todo el epiblasto y trofoblasto, sin embargo queda excluido del endodermo visceral (flecha en F', vista al microscopio de sección del embrión en F en el epiblasto). miR-295 (G) pertenece a la misma familia que miR-291-3p, lo que indica que regulan los mismos genes. Este miRNA se expresa sin embargo con distinto patrón que miR-291-3p. miR-295 se encuentra en estadio E6.5 expresado de forma uniforme tanto en epiblasto como en trofoblasto, pero el nivel de expresión es más elevado en este último linaje.



En conjunto los resultados obtenidos del análisis de los perfiles de expresión de miRNAs tanto en embriones en estadios tempranos de desarrollo como en células madre derivadas del blastocistos nos han permitido conocer qué miRNAs se expresan preferentemente en células TS y trofoblasto embrionario.

miRNAs IMPLICADOS EN LA FUNCIÓN DE CÉLULAS TS

A lo largo de los apartados anteriores hemos descrito la variación en expresión génica en células TS asociada a la falta de función de los miRNAs, centrándonos en último término en aquellos genes cuyo aumento en expresión pueda relacionarse con el fenotipo de diferenciación de las células TS-*Dicer*^{-/-}. Asimismo, hemos obtenido una relación de los miRNAs que se encuentran altamente expresados en células TS y cuya ausencia puede por tanto asociarse al fenotipo de falta de función de *Dicer*.

Uniando estos resultados, en este último apartado trataremos de establecer qué miRNAs pueden tener una mayor importancia a la hora de explicar el fenotipo observado en las células TS-*Dicer*^{-/-}.

7 miRNAs asociados a la variación en expresión génica de las células TS-*Dicer*^{-/-}

Anteriormente hemos visto que la delección de *Dicer* en células TS tiene como resultado la parada de crecimiento y diferenciación de éstas, acompañada de variación en expresión de un número elevado de genes. Aquellos genes cuya expresión aumenta en un primer momento al retirar los miRNAs funcionales serán los principales candidatos a ser regulados por estos represores. Teniendo esto en cuenta, para cada miRNA analizado en el array de PCR cuantitativa buscamos el número de genes diana que presenta dentro de aquellos cuya expresión aumenta en TS a los cuatro días de la infección con AdCre-GFP, utilizando para ello las predicciones dadas por la base de datos TargetScan 4.1 (www.targetscan.org) (Fig. 13A, ver Materiales y Métodos).

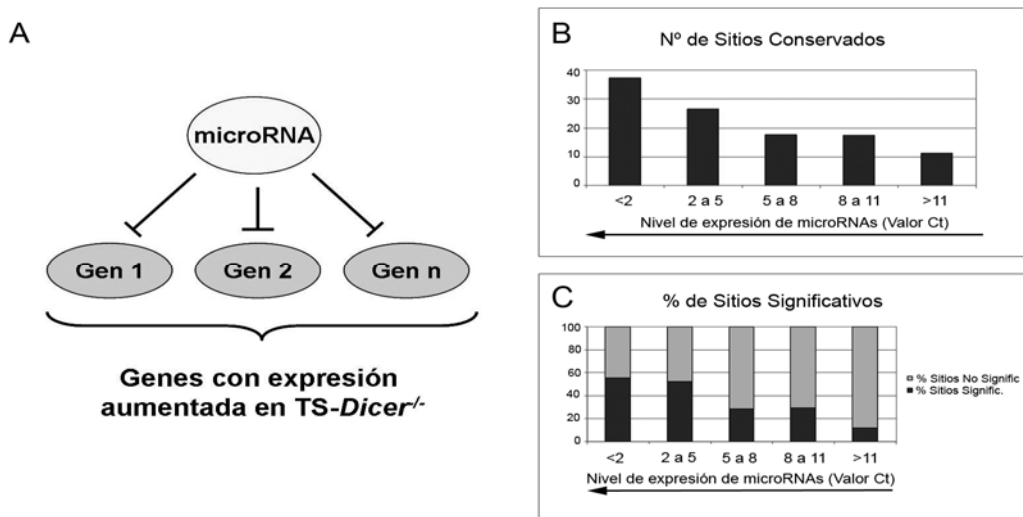


Figura 13. Relación entre el nivel de expresión de miRNAs en células TS y los genes sobre-expresados en las células TS-*Dicer*^{-/-}. Para cada miRNA analizado en el array de PCR cuantitativa buscamos el número de genes diana predichos por la base de datos TargetScan 4.1 (www.targetscan.org) entre aquellos que se sobre-expresan a día cuatro tras la infección con AdCreGFP de células TS-*Dicer*^{fx/fx} (A). El gráfico en (B) representa el número de sitios conservados predichos que cada miRNA tiene entre los genes sobre-expresados, en función a su vez de la expresión normal del miRNA en células TS-*Dicer*^{fx/fx} (medida en valores de Ct). Existe una correlación entre el nivel de expresión del miRNA y el número de genes que regula entre los sobre-expresados en células mutantes, de modo que aquellos más altamente expresados (Ct<5) regulan un mayor número de genes en células TS. El gráfico en (C) representa el porcentaje de miRNAs que poseen mayor número de sitios conservados predichos entre los genes sobre-expresados de los que se esperaría por azar (%Sitios significativos, Chi-Cuadrado con corrección de Yates, p<0.05), dentro de cada categoría de nivel de expresión (dado en valor de Ct). Aquellos miRNAs cuya expresión es elevada en células TS-*Dicer*^{fx/fx} (Ct<5) presentan un número de sitios mayor de lo que se esperaría por azar entre los genes con expresión aumentada al delecionar *Dicer*.

En la figura 13B se representa el número de sitios conservados predichos que cada miRNA tiene entre los genes sobre-expresados en células TS-*Dicer*^{-/-}, en función a su vez del nivel normal de expresión de los miRNAs en las células TS. Aquí podemos ver que existe una correlación entre el nivel de expresión del miRNA en células TS y el número de genes que regula entre los sobre-expresados al delecionar *Dicer*. Esto apunta a los miRNA más altamente

expresados como los principales candidatos a ser responsables del aumento en expresión de muchos de los genes que varían en las células TS-*Dicer*^{-/-}. Al hacer un análisis estadístico de los miRNAs que tienen un número de sitios conservados mayor de los que se esperaría por azar entre los genes sobre-expresados en TS-*Dicer*^{-/-} (estadístico Chi-Cuadrado con corrección de Yates, $p < 0.05$), vemos que una vez más son aquellos miRNAs con expresión elevada en TS los que regulan un mayor número de genes sobre-expresados (Fig. 13C). En ambos casos podemos considerar como límite un valor de Ct de 5, por debajo del cual los miRNAs tienen un nivel de expresión tal que se puede relacionar con una función en células TS.

Tabla 7. miRNAs altamente expresados en TS y con un número significativamente elevado de sitios predichos entre los genes sobre-expresados en las células TS-*Dicer*^{-/-}

Familia	miRNA	Nº Sitios en genes aumentados en TS- <i>Dicer</i> ^{-/-}	Expresión predominante
miR-291-3p/294/295/302/372/373/520	miR-294	36	CM, ET
	miR-295	37	CM, ET
	miR-291a-3p	35	CM, ET
miR-15/16/195/424/497	miR-424	40	TS
	miR-15b	40	Ubicuo
	miR-16	35	Ubicuo
miR-291-5p	miR-291a-5p	27	CM, E6.5, E7.5
	miR-291b-5p	27	CM, ET
miR-25/32/92/363/367	miR-92	36	CM, ET
	miR-25	36	Ubicuo
miR-17-5p/20/93/106/519	miR-20a	72	Ubicuo
	miR-93	73	Células, ET
	miR-106b	72	Ubicuo
	miR-17-5p	72	Ubicuo
	miR-106a	72	ES, ET
miR-19	miR-19b	58	Ubicuo
	miR-19a	58	Ubicuo
miR-26	miR-26a	39	Ubicuo
	miR-26b	39	TS, MEF, ET
miR-30-5p	miR-30b	69	Ubicuo
	miR-30a-5p	70	Células, ET
	miR-30d	70	TS, MEF, E5.5
	miR-30c	69	Ubicuo
miR-467	miR-467b	47	TS, E5.5
miR-27	miR-27a	45	Células
miR-466	miR-466	30	TS, E5.5
miR-322	miR-322	23	TS
miR-374	miR-374-5p	21	ET
miR-125/351	miR-125a	21	Ubicuo
miR-182	miR-182	32	Células ET
miR-130/301	miR-130a	38	Ubicuo
	miR-301	40	TS, Embrión
miR-34/449	miR-34a	24	Células
	miR-699	23	Ubicuo
miR-29	miR-29b	35	TS
	miR-29a	33	TS
	miR-29c	34	TS
let-7/98	let-7a	36	TS, MEF, E5.5
	let-7g	36	TS, MEF, E5.5

Índice: CM-células madre (ES y TS); ET-embrión temprano; Células-TS, ES y MEF; Embrión: E5.5 a E8.5

En la tabla 7 se muestran, agrupados en familias, todos los miRNAs altamente expresados (valor de Ct<5) en células TS-*Dicer*^{fx/fx}, así como su patrón de expresión y el número de genes que se ha predicho que regulan entre los sobre-expresados en TS-*Dicer*^{-/-}. Tanto en la figura 13 como en la tabla 7 se representa únicamente el número de sitios conservados, ya que son en

principio los que presentan mayor fiabilidad en la predicción. Sin embargo en muchos casos los sitios no conservados son también funcionales (Baek et al., 2008; Grimson et al., 2007). En este sentido, el mismo análisis descrito anteriormente se llevó a cabo para los sitios no conservados, obteniendo resultados muy similares, en los que los miRNAs con un valor de expresión de Ct inferior a 5 son aquellos que regulan un mayor número de genes entre los sobre-expresados en células TS-*Dicer*^{-/-} (datos no mostrados).

8 miRNAs relacionados con el fenotipo de las células TS-*Dicer*^{-/-}

La asociación de los datos de expresión de miRNAs con los datos de variación génica en células TS-*Dicer*^{-/-} apuntan a los miRNA altamente expresados (valor de Ct<5) como principales candidatos a ser responsables del fenotipo observado. Por otro lado, los datos de expresión de mRNA en células mutantes indican la existencia de diferentes rutas de señalización o factores cuya alteración puede ser responsable de la diferenciación (ver apartado 4). Si analizamos qué miRNAs altamente expresados se encuentran como potenciales reguladores de los genes implicados en estas vías tendremos una idea de qué rutas pueden estarse viendo alteradas directamente por la falta de estos represores. La tabla 8 muestra los resultados de este análisis

Tabla 8. Relación entre vías afectadas y miRNAs que las regulan

Gen	Cambio Día 4	N° Sitios para miRNAs con Ct<5	
		Conservados	No conservados
<i>Dusp9</i>	1.5	0	1
<i>Rasa2</i>	1.4	0	0
<i>Cdkn1c</i>	1.5	2	1
<i>Cdk6</i>	1.9	11	14
<i>Cdkn1a</i>	1.9	5	13
<i>Lats2</i>	1.3	17	6
<i>Rbl2</i>	1.3	10	6
<i>Lif</i>	1.5	10	10
<i>Il6st</i>	2.0	0	0
<i>Col4A1</i>	1.4	6	0
<i>Col4A2</i>	1.3	5	8

La inhibición de la vía de Fgf4 por parte de *Dusp9* y/o *Rasa2* podría ser en parte responsable de la diferenciación de las células mutantes, sin embargo de entre los miRNAs expresados en TS hay predicho únicamente uno cuya diana sea *Dusp9*, mientras que *Rasa2* no está regulado por ningún miRNA altamente expresado en TS. Así pues descartamos la alteración de esta vía como un efecto directo de la ausencia de miRNAs. En el caso de la vía de Lif parece que esta propia citoquina podría estar en TS directamente regulada por miRNAs, pero no así su co-receptor Il6st, que no posee ningún sitio de unión para miRNAs altamente expresados en TS. La contribución a la diferenciación por sobre-expresión de *Col4a* podría constituir también un efecto directo de la retirada de miRNAs dado el número elevado de sitios predichos que posee este gen. Por último, el número de sitios para miRNAs altamente expresados en TS que poseen los genes implicados en ciclo celular es muy elevado, por lo que la desregulación de éstos podría estar directamente relacionada con la ausencia de miRNAs, lo que conllevaría

tanto a diferenciación (debida a un aumento de *Cdkn1c*, *Cdkn1a* y *Cdk6*) como a la parada de crecimiento (debida a un aumento de *Cdkn1a*, *Lats2* y *Rbl2*). Así pues seleccionamos como principales candidatos ser responsables del fenotipo debido directamente a la ausencia de miRNAs los genes: *Cdkn1c*, *Cdk6*, *Cdkn1a*, *Lats2*, *Rbl2*, *Lif* y *Col4a*.

Tabla 9. Familias de miRNAs que regulan a los genes candidatos

<i>Cdkn1c</i>	<i>Cdk6</i>	<i>Cdkn1a</i>	<i>Lats2</i>	<i>Rbl2</i>	<i>Lif</i>	<i>Col4a</i>
miR-92	miR-291-3p	miR-291-3p	miR-291-3p	miR-291-3p	miR-291-3p	miR-291-3p
miR-466	miR-424	miR-424	miR-424	miR-424	miR-424	miR-29
	miR-29	miR-29	miR-467b	miR-17-5p	miR-29	miR-467b
	miR-467b	miR-467b	miR-92		miR-17-5p	miR-17-5p
	miR-466	miR-17-5p	miR-182		miR-291-5p	let-7
	miR-182		miR-30-5p		miR-125	
	miR-26		miR-27		miR-130	
	miR-34a		miR-290			
			miR-374			

Los miRNAs pertenecientes a una misma familia comparten los 7-8 nucleótidos que constituyen el sitio de unión principal al RNA mensajero diana y que va a determinar la mayor parte de la especificidad de la unión. De este modo, aquellos miRNAs que se encuentren en la misma familia regularan potencialmente los mismos genes (Chen and Rajewsky, 2007; Grimson et al., 2007). Gran parte de los miRNAs altamente expresados en TS se agrupan en familias con gran número de miembros, muchos de ellos expresados en niveles altos no sólo en TS sino también en otros tipos celulares o en embriones (ver tabla 7). Esto indica la importancia que estas familias tienen en la regulación génica. En la tabla 9 se muestran las familias de miRNA que están predichas como reguladoras de los genes candidatos seleccionados (en cada caso el miRNA más representativo es el que da nombre a la familia, el resto de miembros se encuentran listados en la tabla 7). Hay 7 familias especialmente relevantes (resaltadas en el mismo color en la tabla 9) ya que regulan varios de los genes candidatos cada una y además entre las ocho cubren la regulación de los siete genes. De estas familias, miR-29, miR-467 y miR-466 se encuentran expresadas de forma predominante en células TS (miR-29) o en TS y embrión temprano (miR-467 y miR-466). La familia miR-291-3p es específica de linajes tempranos (TS, ES y embrión temprano) y por último los miembros de las familias miR-424, miR-17-5p y miR-92 se encuentran expresados de forma más o menos ubicua (para expresión específica de cada miembro ver tabla 7).

Así pues, son varios los genes que pueden considerarse como candidatos a ser regulados directamente por miRNAs y cuya desregulación debida a la falta de éstos puede contribuir a la diferenciación y parada del crecimiento de las células TS. Además, existen familias de miRNAs expresadas tanto de forma específica como ubicua que parecen tener un mayor peso sobre la regulación génica de células TS.

Discusión

Los eventos que tienen lugar durante el desarrollo embrionario temprano son determinantes para la correcta formación de cualquier organismo. La aparición de un desarrollo completamente intrauterino en el embrión de mamíferos va acompañada de la formación de estructuras extraembrionarias altamente especializadas, que contribuyen no sólo al soporte nutricional del embrión sino también a su correcta especificación de patrón en las etapas tempranas. Así pues, tanto tejidos embrionarios como extraembrionarios han de ser considerados a la hora de estudiar los procesos que tienen lugar durante estos primeros momentos del desarrollo en mamíferos.

En este trabajo hemos abordado el estudio de la regulación post-transcripcional mediada por miRNAs durante las primeras fases de formación del trofoblasto que dará lugar a la placenta del ratón. Para ello hemos utilizado como modelo el mutante de *Dicer*, una proteína esencial para la maduración de los miRNAs cuya delección ha servido para abordar el estudio de la función de estos represores durante el desarrollo de numerosos órganos y tejidos (Harfe et al., 2005; Harris et al., 2006; O'Rourke et al., 2007). El análisis tanto *in vivo* como *in vitro* de la falta de función de *Dicer* en el trofoblasto nos ha permitido mostrar cómo los miRNAs se requieren en este linaje para el mantenimiento de su población de células madre en condiciones de multipotencia y auto-renovación.

1 Los miRNAs se requieren para mantener las células TS en el embrión

En el blastocisto de ratón a los 3.5 días de desarrollo encontramos células madre para cada uno de los primeros linajes especificados: células ES precursoras del embrión, células TS precursoras del trofoblasto y células XEN precursoras del endodermo extraembrionario. A lo largo del desarrollo estas células irán restringiendo su multipotencia y por tanto su capacidad de diferenciación en los diferentes tipos celulares que conforman cada uno de los linajes de los que son precursoras (Rossant and Tam, 2009). Las células TS mantienen su capacidad de auto-renovación y su multipotencia durante un periodo de tiempo comprendido entre los 3.5 días del desarrollo –el momento de su especificación– hasta aproximadamente los 7.5 días de desarrollo (Tanaka et al., 1998).

El ectodermo extraembrionario (ExE) a los 6.5 días de desarrollo está compuesto por poblaciones celulares con distinto grado de diferenciación, y entre ellas se encuentran las células TS, que son las que se mantienen como precursoras de trofoblasto indiferenciadas. Aquellos genes normalmente expresados en todo el territorio del ExE, como son *Fgfr2* o *Rhox5* mantienen sus niveles de expresión normales en los embriones mutantes, mientras que la drástica bajada en expresión de genes necesarios para el mantenimiento de las células TS como son *Cdx2*, *Esrrb* o *Eomes* muestra cómo esta población celular se encuentra comprometida en ausencia de miRNAs. El defecto específico de este tipo celular se pone también de manifiesto cuando se examina la expresión de genes propios de poblaciones diferenciadas, que presentan niveles normales de expresión en embriones mutantes. Así pues,

los resultados obtenidos del análisis de embriones *Dicer*^{-/-} indican que a los 6.5 días de desarrollo la única población celular afectada por la ausencia de miRNAs en el trofoblasto son las células TS. El defecto en el mantenimiento de la población de precursores explicaría así la notable reducción de tamaño en la región del ectodermo extraembrionario que se observa a los 7.5 días de desarrollo.

2 Papel de los miRNAs en la proliferación de las células TS

El mantenimiento de las células TS como precursoras dentro del ectodermo extraembrionario requiere de su capacidad de proliferar en ausencia de diferenciación. La proliferación en el ectodermo extraembrionario se encuentra disminuida en los embriones *Dicer*^{-/-} sin que exista un aumento de la muerte celular. La bajada de proliferación en este tejido estará afectando presumiblemente a la capacidad de auto-renovación de las células TS, lo que tiene como consecuencia la disminución en el número de precursores acompañada de bajada en expresión de genes propios de este tipo celular.

Este defecto de proliferación ligado a la falta de miRNAs se pone de manifiesto cuando se generan células TS mutantes para *Dicer* en cultivo. El sistema de delección por infección utilizado en este trabajo permite detectar los primeros efectos fenotípicos que se producen tras la retirada de los miRNAs. En este sentido, la delección de *Dicer* tiene como consecuencia la bajada drástica de proliferación de las células TS, notable a partir de los 4-5 días de la infección con el adenovirus AdCre-GFP, cuando la delección de *Dicer* ya ha tenido lugar y los niveles de miRNAs maduros probablemente han disminuido de forma importante.

Las poblaciones de células madre se caracterizan por su capacidad de división durante largos periodos de tiempo, para lo cual han de regular de manera especial los mecanismos de control del ciclo celular. Los miRNAs se han implicado en la regulación de la transición de G1 a S del ciclo celular en células madre de distintos organismos, como son las células germinales de *Drosophila* o las células madre embrionarias (ES) de ratón. Así, la falta de miRNAs en estas poblaciones celulares trae consigo un aumento en expresión de inhibidores del ciclo celular como son p21 en ratón o Dacapo (homólogo a p21/p27) en la mosca, provocando una bajada de proliferación (Hatfield et al., 2005; Wang et al., 2008; Yu et al., 2009).

El concreto en células ES, la delección de genes codificantes para las proteínas Dgcr8 o *Dicer* tiene como consecuencia la disminución en la tasa de proliferación acompañada de la acumulación de células en fase G1 (Kanellopoulou et al., 2005; Murchison et al., 2005; Wang et al., 2007). Este defecto en las células ES mutantes para *Dicer* parece ser más acusado en el momento de su derivación, de modo que al cabo de cierto tiempo recuperan la capacidad proliferativa, probablemente debido a la acumulación de mutaciones o de cambios estables en la expresión génica que compensan el defecto inicial (Murchison et al., 2005). En cambio en el caso de la delección de *Dgcr8*, las células ES mutantes presentan un defecto de proliferación más estable, que no se rescata con el tiempo (Wang et al., 2007). Este último modelo ha sido

el utilizado para el estudio de la regulación de la proliferación por miRNAs en células ES (Wang et al., 2008). Wang y colaboradores llevaron a cabo una búsqueda sistemática de miRNAs que pudieran estar implicados en el defecto de proliferación de las células ES mutantes para *Dgcr8*. Mediante la re-introducción uno a uno de miRNAs en las células mutantes detectaron 14 miRNAs capaces de rescatar el fenotipo de proliferación. La mayor parte de ellos pertenecen a dos familias altamente expresadas en células ES y cuya región semilla es prácticamente idéntica. Estas son las familias a las que pertenecen miR-291-3p y miR-17-5p. En este trabajo relacionan cinco de los miRNAs identificados (miR-291-3p, miR-291b-3p, miR-294, miR-295 y miR-302) con la represión de los inhibidores del paso de fase G1 a S del ciclo celular *Cdkn1a* (p21), *Lats2* y *Rbl2*. La ausencia de estos miRNAs, a los que denominan ESCC (del inglés, ES cell-specific cell cycle-regulating miRNAs), sería por tanto la responsable del aumento en los niveles de inhibidores del ciclo celular, provocando una bajada de proliferación en las células ES (Wang et al., 2008).

Se desconoce el mecanismo exacto por el cual la ausencia de miRNAs en las células TS mutantes para *Dicer* provoca una bajada de la proliferación, sin embargo son varios los indicios que nos hacen pensar que pudiera ser similar al implicado en el defecto descrito para las células ES mutantes para *Dgcr8*. Así, es notable el aumento en células TS-*Dicer*^{-/-} de la expresión de los genes codificantes para los inhibidores del paso de fase G1 a S del ciclo celular p21, Rbl2 y Lats2 del mismo modo que sucede en células ES mutantes para *Dgcr8*. Por otro lado, entre los miRNAs más expresados en las células TS encontramos miembros de las dos familias que se han implicado en el control de la proliferación de las células ES: las familias de las que forman parte miR-17-5p y miR-291-3p (Fig. 1)

Cuatro de los cinco miembros de la familia miR-17-5p/20/93/106 se encuentran altamente expresados en células TS y tres de ellos (miR-20a, miR-106b y miR-17-5p) han sido considerados como ubicuos en este trabajo, puesto que su nivel de expresión es muy elevado en todas las muestras celulares y embrionarias analizadas. Esta familia de miRNAs y en particular los miembros integrantes del cluster miR-17-92 y de sus clusters parálogos se encuentran implicados en las redes que controlan la proliferación y la muerte celular tanto durante el desarrollo normal como en procesos oncogénicos (Mendell, 2008). La sobre-expresión del cluster miR-17-92 se ha asociado a distintos tipos de tumores, poniendo de manifiesto la naturaleza de oncogén que este locus tiene (Zhang et al., 2007). En el desarrollo, se ha visto que la sobre-expresión del cluster miR-17-92 en pulmón provoca un aumento notable de la proliferación a la vez que se inhibe la diferenciación celular (Lu et al., 2007). Por el contrario, los ratones mutantes para este cluster presentan hipoplasia pulmonar (Ventura et al., 2008). Entre los genes cuya regulación por miembros de la familia miR-17-5p se ha comprobado experimentalmente se encuentran *Cdkn1a* (p21) y *Rbl2* (Ivanovska et al., 2008; Lu et al., 2007), que como ya hemos mencionado se encuentran sobre-expresados en las células TS mutantes para *Dicer*.

En cuanto a la familia miR-291-3p/294/295/302, tres de sus miembros se encuentran altamente expresados en células TS (miR-291-3p, miR-294 y miR-295). La ausencia de los miembros de esta familia puede estar contribuyendo asimismo a la sobre-expresión de los inhibidores de progresión a través del ciclo celular p21, Rbl2 y Lats2 del mismo modo en que lo hacen en las células ES mutantes para *Dgcr8*. Varios miembros de esta familia de miRNAs pertenecen al cluster-290, que ha sido descrito como específico de células ES, puesto que se encuentra altamente expresado en este tipo celular y su expresión decae con la diferenciación (Houbaviy et al., 2003). Nuestros resultados muestran que este cluster, y por tanto los miembros de la familia miR-291-3p que a él pertenecen, se encuentra altamente expresado en células madre derivadas del blastocisto (ES y TS), así como en el embrión temprano en post-implantación, disminuyendo su expresión a medida que avanzamos en el desarrollo. Así pues, al contrario de lo que sucede con la familia miR-17-5p, cuya expresión ubicua podría contribuir al control del ciclo celular de forma más generalizada, la familia miR-291-3p tendría un papel específico sobre el control de la proliferación durante el desarrollo embrionario temprano.

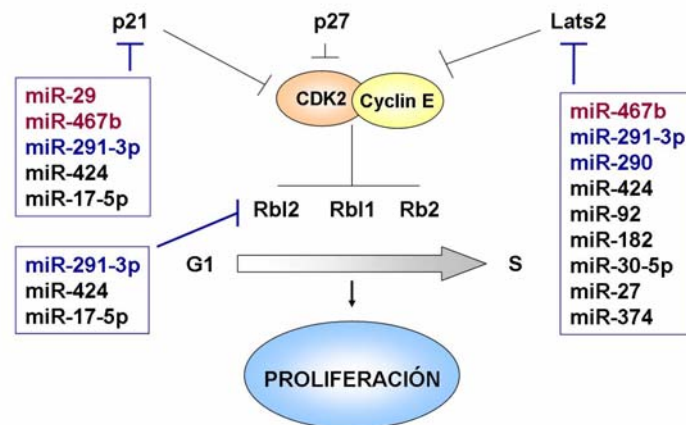


Figura 1. Implicación de los miRNAs en el control de la proliferación de células TS. Mediante la represión de factores implicados en el bloqueo del paso de fase G1 a fase S del ciclo celular como son p21, Lats2 o Rbl2, los miRNAs pertenecientes a las familias representadas estarían contribuyendo a promover la proliferación de las células TS. En este mecanismo, ya descrito previamente para células ES, intervienen miRNAs cuya expresión es fundamentalmente ubicua (en negro) o específica de células madre derivadas del blastocisto (ES y TS, resaltados en azul) o en células TS (en rojo).

Aunque el modo de actuación descrito para ambas familias de miRNAs es el mismo -la represión de inhibidores de la progresión de fase G1 a S del ciclo celular- los mecanismos por los cuales los miRNAs intervienen en la regulación de la proliferación son presumiblemente más complejos y muy probablemente requieren de la intervención de otras familias. No obstante, la importancia de las familias miR-17-5p y miR-291-3p durante el desarrollo embrionario temprano así como en el mantenimiento tanto de células ES como TS se refleja en el hecho de que ambas tienen varios miembros que se encuentran entre los más expresados en estos estadios y tipos celulares (Tabla 6 de Resultados, (Calabrese et al., 2007)). Además,

en el caso concreto de las células TS cada uno de los miembros altamente expresados de estas familias podría regular potencialmente un número elevado de genes entre los sobre-expresados en las células TS-*Dicer*^{-/-}.

3 Papel de los miRNAs en el mantenimiento de la multipotencia de las células TS

El mantenimiento de una población de progenitores dentro de un determinado tejido depende de su capacidad para preservar las condiciones de auto-renovación y multipotencia. Para el mantenimiento de la auto-renovación se precisa un control fino de la proliferación, y como hemos visto anteriormente ésta se encuentra regulada en cierta medida por miRNAs en las células TS al igual que se había descrito previamente para las células ES. En cuanto la multipotencia, esta propiedad supone para las células el mantenerse en un estado indiferenciado pero ser capaces a su vez de diferenciarse en un momento dado tomando un destino determinado en respuesta a estímulos concretos.

Como ya se ha mencionado previamente, los embriones mutantes para *Dicer* presentan un defecto específico en el mantenimiento de la población de células multipotentes dentro del ectodermo extraembrionario. Cuando analizamos los marcadores de diferenciación vemos que éstos no se encuentran alterados. Así, *Asc/2* se mantiene expresado en la zona del cono ectoplacentario (EPC) en estadio E6.5 y aun se expresa en niveles normales en estadio E7.5, cuando el ectodermo extraembrionario se ha reducido considerablemente en la mayor parte de los embriones de modo que únicamente queda la zona de diferenciación del EPC. Del mismo modo, *Rhox5* mantiene su expresión en la zona de diferenciación del EPC tanto en estadio E6.5 como en E7.5. En cultivo las células TS son incapaces de mantener la multipotencia en ausencia de miRNAs, de modo que comienzan a diferenciarse entre los 3 y 4 días tras la delección de *Dicer*. Este proceso de diferenciación se refleja en la bajada de expresión de genes propios del estado multipotente y la subida en la expresión de genes propios de tipos celulares diferenciados.

Los mecanismos destritos para la diferenciación de las células TS en cultivo incluyen la falta de señalización por Fgf4 a través de la activación de Erk1/2 y/o por factores de crecimiento de la familia TGFbeta a través de la activación de Smad2/4, que son del mismo modo necesarios para el mantenimiento de la población de células TS dentro del ectodermo extraembrionario *in vivo* (Erlebacher et al., 2004; Guzman-Ayala et al., 2004; Rossant and Cross, 2001; Tanaka et al., 1998). Las células TS-*Dicer*^{-/-} presentan sobre-expresión de dos inhibidores de la vía de Fgf: *Rasa2* y *Dusp9*. *Rasa2* es un inhibidor de la proteína Ras, que forma parte de la vía de activación de Erk1/2 (Lu et al., 2008; Yarwood et al., 2006), mientras que *Dusp9* es una fosfatasa expresada en placenta que de-fosforila a Erk1/2 inactivándolo (Christie et al., 2005; Patterson et al., 2009). Además existe una ligera disminución en la expresión del receptor de Fgf4, *Fgfr2*. Todo esto podría significar una bajada en la señalización por Fgf4 que resultaría en la diferenciación de las células TS. Sin embargo son varios los motivos que nos incitan a

descartar la alteración en esta vía como causa de la diferenciación por ausencia de miRNAs. En primer lugar, no parece que en el embrión mutante se aprecie una disminución en la expresión de los genes implicados en la señalización por Fgf4, de modo que la expresión de *Fgf4* en el epiblasto es normal, al igual que sucede con la de su receptor *Fgfr2* en el trofoblasto. Además, resultados obtenidos en el laboratorio del Dr. T. Rodríguez indican una correcta activación de Erk1/2 en el trofoblasto de los embriones mutantes para *Dicer* en estadio E6.5 (comunicación personal). Por otro lado, ninguno de los miRNAs altamente expresados en células TS presenta sitios de unión predichos en la región 3'UTR del gen *Rasa2*, y para *Dusp9* únicamente hay un sitio de unión no conservado predicho para miRNAs altamente expresados en TS. Por esta razón, aunque la señalización por Fgf4 a través de la vía de Erk1/2 se encontrase alterada en las células TS-*Dicer*^{-/-}, resulta difícil atribuir esta alteración a un efecto directo de la falta de miRNAs. Es notable además el hecho de que la retirada de Fgf4 únicamente provoca la diferenciación de las células TS hacia células gigantes a los seis o siete días en condiciones normales (Hughes et al., 2004; Tanaka et al., 1998), mientras que el efecto de diferenciación que observamos tras la retirada de los miRNAs se produce mucho antes; esto indicaría que la causa primera de la diferenciación tras la delección de *Dicer* es probablemente independiente de la señalización por Fgf4.

En cuanto a la señalización por factores de crecimiento de la familia TGFbeta, al igual que sucedía en el caso de Fgf4 ninguno de los genes implicados en esta vía parece tener expresión alterada en el embrión. La expresión de *Nodal* en el epiblasto de los embriones mutantes es normal, así como la de las convertasas *Spc1* y *Spc4* en el trofoblasto. Asimismo, dado que la señalización por Nodal se requiere para la especificación del mesodermo en la región posterior del embrión (Tam and Loebel, 2007), la correcta expresión del gen propio de mesodermo *Eomes* en esta región indicaría el buen funcionamiento de la vía en los embriones mutantes. Por otro lado, no hemos encontrado ningún inhibidor de la vía sobre-expresado en las células TS tras la retirada de los miRNAs, por lo que la inhibición de la señalización por factores de la familia TGFbeta no parece ser un mecanismo implicado en la diferenciación de las células TS en ausencia de miRNAs.

No obstante, hay que tener en cuenta que el mecanismo de acción de los miRNAs más generalizado en células animales es la represión a nivel de traducción, que aunque en la mayor parte de los casos acaba afectando también a los niveles de mRNA, la disminución en los niveles de proteína será siempre mayor (Baek et al., 2008; Filipowicz et al., 2008; Selbach et al., 2008). Por este motivo no podemos descartar que haya otros inhibidores de estas dos vías que se estén viendo afectados únicamente en los niveles de proteína, lo que no se reflejaría en una variación en el array de mRNA.

4 Control del ciclo celular de las células TS por miRNAs

La falta de señalización por Fgf4 en células TS tiene como consecuencia la diferenciación de éstas en los distintos tipos celulares que encontramos en el linaje del trofoblasto (Hughes et al., 2004). Uno de estos tipos celulares, las células gigantes, se caracteriza por la endo-reduplicación de su DNA, que es un fenómeno que raramente se da de forma natural durante el desarrollo en las células de mamíferos. La endo-reduplicación del DNA supone la inhibición controlada de la división celular tras rondas sucesivas de síntesis, evitando a su vez que las células mueran por apoptosis en respuesta a este fenómeno. Recientemente se ha descrito un mecanismo por el cual la falta de señalización por Fgf4 puede provocar la endo-reduplicación y diferenciación celular hacia células gigantes (Ullah et al., 2008). En este caso, la falta del factor de crecimiento provoca un aumento de los niveles de los inhibidores de kinasa dependiente de ciclina Cdkn1c (p57) y Cdkn1a (p21). Por un lado p57 inhibe a la proteína kinasa dependiente de ciclina Cdk1 implicada en la entrada y progreso a través de la fase de mitosis; y por otro lado p21 reprime a la kinasa de control Chk1, inhibiendo los mecanismos de respuesta frente al daño en el DNA desencadenados como consecuencia de la endo-reduplicación y permitiendo así la supervivencia. La activación de este mecanismo de diferenciación hacia células gigantes de trofoblasto supone el situar a las células en un estado de síntesis continua acompañado de una inhibición de la mitosis (Ullah et al., 2008). Como hemos mencionado anteriormente, no parece que exista una falta de activación de la vía de Erk1/2 en respuesta a Fgf4 en las células TS mutantes para *Dicer*, sin embargo sí se produce un aumento en los niveles de expresión de los dos inhibidores de kinasa situados bajo la regulación de Erk1/2, p57 y p21. Asimismo, parece que las células TS responden a la retirada de los miRNAs de manera similar a como lo hacen en el mecanismo de diferenciación hacia células gigantes regulado por p57, esto es, disminuyendo los niveles de *Cdk1* y *Chk1* de forma que se permita la endo-reduplicación del DNA y la supervivencia celular (Fig. 2).

Para que las células progresen de la fase G1 a S durante la endo-reduplicación se habrá de reprimir en cierto modo a los inhibidores este paso, como son las proteínas de la familia de retinoblastoma (Malumbres and Barbacid, 2009). En este sentido es notable el aumento de la expresión de la proteína kinasa dependiente de ciclina Cdk6 en las células TS tras la delección de *Dicer*. Cdk6 forma complejos con las ciclinas de tipo D provocando una represión parcial de las proteínas de la familia del retinoblastoma. Aunque en trabajos previos no se ha implicado a Cdk6 en el mecanismo de diferenciación de células TS, su aumento en expresión podría contribuir en las células TS-*Dicer*^{-/-} a la progresión de fase G1 a S del ciclo celular necesaria para que se de la endo-reduplicación. En este sentido, se ha descrito la regulación de *Cdk6* por distintos miRNAs, entre los que se encuentran los miembros de la familia miR-16, expresada en niveles altos en células TS (Linsley et al., 2007; Pierson et al., 2008; Sun et al., 2008).

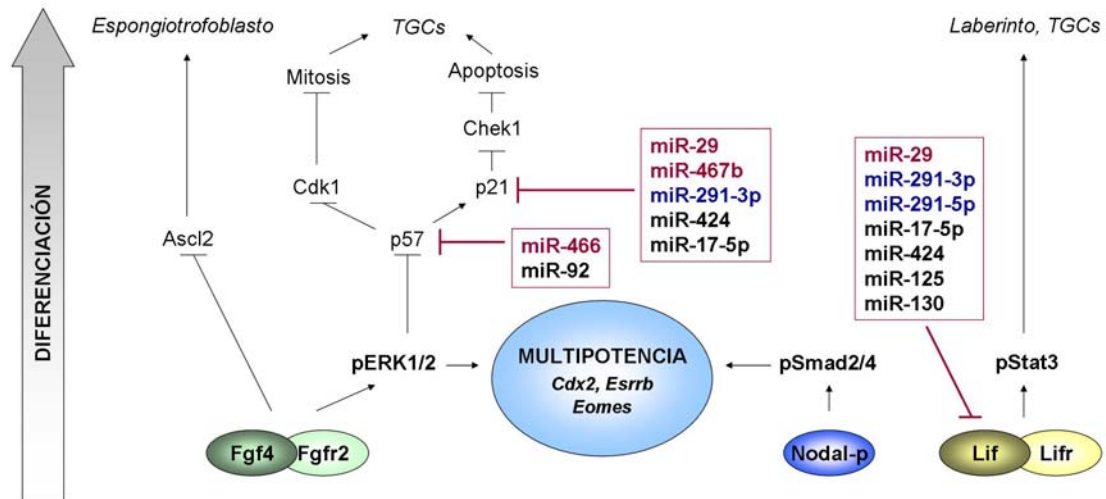


Figura 2. Implicación de los miRNAs en el mantenimiento de la multipotencia de células TS. El esquema muestra las principales vías de señalización que hasta el momento se han implicado en el mantenimiento de las células TS *in vivo* e *in vitro*; asimismo, se muestran las familias de miRNAs que según nuestros resultados constituyen buenos candidatos a regular distintos factores implicados en estas vías. La regulación por miRNAs estaría contribuyendo al mantenimiento de las células TS indiferenciadas actuando a distintos niveles. Entre ellos estaría la contribución al control de la endo-reduplicación mediada por p57 que tiene lugar durante la diferenciación hacia células gigantes de trofoblasto (TGCs) en ausencia de Fgf4. Otro mecanismo de diferenciación descrito aunque no bien caracterizado es el mediado por la activación de Stat3 en respuesta a la señalización por Lif. Este podría del mismo modo estar regulado en parte por miRNAs, contribuyendo por tanto al mantenimiento de las células TS indiferenciadas. Las familias de miRNAs representadas en rojo tienen expresión enriquecida en células TS; las familias en azul representan miRNAs altamente expresados en células ES y TS; y las familias en negro se encuentran expresadas de forma ubicua.

Así como no encontramos miRNAs candidatos a regular la vía de Erk1/2 en células TS, son varias las familias de miRNAs altamente expresadas en las células TS que podrían regular tanto a p57 como a p21 y Cdk6. Es destacable el hecho de que entre los miRNAs que poseen sitios de unión en la región 3'UTR de estos tres genes se encuentran aquellos pertenecientes a familias expresadas de forma predominante tanto en células TS como en el trofoblasto *in vivo*, como son miR-466 o miR-467b. El fenómeno de endo-reduplicación programada asociado a la diferenciación durante el desarrollo normal únicamente se da en dos tipos celulares en mamíferos, las células gigantes de trofoblasto en la placenta y los megacarioblastos en la médula ósea (Ravid et al., 2002). Por otro en las células ES la inhibición de Cdk1 conlleva la muerte celular por apoptosis (Ullah et al., 2008). En este sentido ciertas familias de miRNAs mayoritariamente expresadas en células TS y trofoblasto podrían estar regulando aquí la diferenciación mediante el control de un mecanismo muy específico de este linaje (Fig. 2).

La falta de regulación de la endo-reduplicación por miRNAs explicaría la diferenciación de las células TS hacia células gigantes de trofoblasto en ausencia de Dicer. Sin embargo, el análisis de los genes sobre-expresados tras la delección de *Dicer* indica que en las células mutantes se produce también diferenciación hacia otros tipos celulares. Deben existir por tanto otros mecanismos independientes del control de p57 que actúan en las células mutantes. Entre los factores que podrían estar contribuyendo a la diferenciación se encuentra la sobre-expresión

de los genes codificantes para la citoquina *Lif*, cuya señalización a través de *Stat3* se ha relacionado con la diferenciación de células TS (Fig. 2) (Takahashi et al., 2003; Takahashi et al., 2008); o para dos subtipos de el colágeno IV (*Col4A1* y *Col4A2*), capaces de provocar cierta transdiferenciación de células ES a TS induciendo la expresión entre otros de genes propios de células TS diferenciadas (Schenke-Layland et al., 2007). Tanto *Lif* como *Col4A1* y *Col4A2* tienen un número elevado de sitios de unión para miRNAs altamente expresados en células TS en sus regiones 3'UTR.

5 Diferencias en el control de la diferenciación por miRNAs en células ES y TS

Al igual que las células TS, las células ES tienen la multipotencia comprometida en ausencia de miRNAs, pero en este caso el fenotipo que se obtiene al delecionar tanto *Dicer* como *Dgcr8* es el contrario: las células ES mutantes para cualquiera de estas dos proteínas son incapaces de diferenciarse. En ambos casos parece que el defecto va asociado a la incapacidad para silenciar genes de pluripotencia, como *Pou5f1* (*Oct4*), cuando se sitúan en condiciones de diferenciación o se induce la formación de “cuerpos embrionarios” (agregados de células que diferencian las tres capas germinales) (Kanellopoulou et al., 2005; Wang et al., 2007). Se ha relacionado de nuevo a la familia de miR-291-3p con el defecto de diferenciación de las células ES mutantes para *Dicer*. En este caso la sobre-expresión de *Rbl2* debido la ausencia de su represor miR-291-3p conllevaría la inhibición de las metiltransferasas *Dnmt1*, *Dnmt3a* y *Dnmt3b* encargadas del silenciamiento, entre otros, de *Pou5f1* durante la diferenciación (Sinkkonen et al., 2008). Sin embargo la reintroducción de miR-291-3p únicamente rescata la falta de represión de *Pou5f1* (*Oct4*) en las células ES mutantes para *Dicer* pero no en las mutantes para *Dgcr8* (Wang et al., 2008). Esta diferencia podría radicar en el hecho de que para derivar las líneas estables de células ES mutantes para *Dicer* se requiere un número elevado de pases en cultivo, por lo que se pueden acumular modificaciones estables en la expresión génica que enmascaren el efecto inicial que los miRNAs tienen sobre la diferenciación (Murchison et al., 2005). En este sentido el abordaje utilizado este trabajo para la generación de células TS mutantes para *Dicer* permite detectar la variación génica que aparece de forma más inmediata tras la retirada de los miRNAs maduros, de forma que no parece probable que el defecto observado en las células TS mutantes a los cuatro días de la infección con AdCre-GFP pueda deberse a modificaciones epigenéticas.

Por otro lado, al contrario de lo que sucede cuando se compara el fenotipo de células ES mutantes para *Dicer* con el epiblasto embrionario mutante, el fenotipo observado en las células TS mutantes para *Dicer* obtenidas de esta forma se asemeja en gran medida al observado *in vivo* en embriones *Dicer*^{-/-}. Asimismo, el perfil de expresión de miRNAs obtenido para las células TS nos ha permitido detectar miRNAs expresados específicamente en el trofoblasto embrionario. Esto refleja el hecho de que la población de células TS presente en el trofoblasto del embrión en estadios previos a la gastrulación es equiparable en gran medida a la población

que encontramos en el blastocisto, y pone de manifiesto la gran utilidad del uso de células TS en cultivo como aproximación al estudio de los procesos que tienen lugar en el trofoblasto *in vivo*.

Es evidente que durante el desarrollo temprano del ratón la acción de los miRNAs va encaminada a la regulación de los procesos de auto-renovación y multipotencia de las células madre de los distintos linajes embrionario y extraembrionario. La auto-renovación estaría regulada en ambos casos por miRNAs que cooperan en el control del ciclo celular mediante mecanismos comunes a otros muchos tipos celulares (es el caso de la familia miR-17-5p) y propios de células madre del embrión temprano (controlados en este caso por miembros de la familia miR-291-3p). La regulación de la multipotencia mediada por miRNAs es sin embargo diferente en las células madre embrionarias y extraembrionarias. Así, en células ES la familia miR-291-3p permite la diferenciación de éstas mediante el silenciamiento de genes de pluripotencia. Por el contrario en células TS la regulación por distintas familias de miRNAs entre las que se encuentran algunas específicas de este tipo celular como son miR-466 o miR-467b, podría estar reprimiendo la diferenciación hacia células gigantes mediante el control de la endo-reduplicación. No obstante, este mecanismo no explica toda la diferenciación que se observa en las células TS mutantes para *Dicer*, por lo que quedarían por determinar otros mecanismos de diferenciación controlados por miRNA en este tipo celular que contribuyan al fenotipo.

El análisis del fenotipo de células TS mutantes de *Dicer* combinado con la caracterización del perfil de expresión de miRNAs en este tipo celular nos ha permitido seleccionar miRNAs que constituyen buenos candidatos a la regulación de la proliferación y multipotencia en células TS. Estos resultados servirán de base para posteriores estudios funcionales con miRNAs concretos.

6 Función de los miRNAs en el desarrollo temprano del embrión de ratón

El papel fundamental de los miRNAs durante el desarrollo embrionario temprano se hace evidente cuando se generan mutantes carentes de estos reguladores, que en todos los casos tienen un fenotipo de letalidad en estadios de pre-gastrulación (Bernstein et al., 2003; Morita et al., 2007; Wang et al., 2007). En la figura 3 se resume el fenotipo que hemos observado en este trabajo para los embriones y células madre mutantes para *Dicer*.

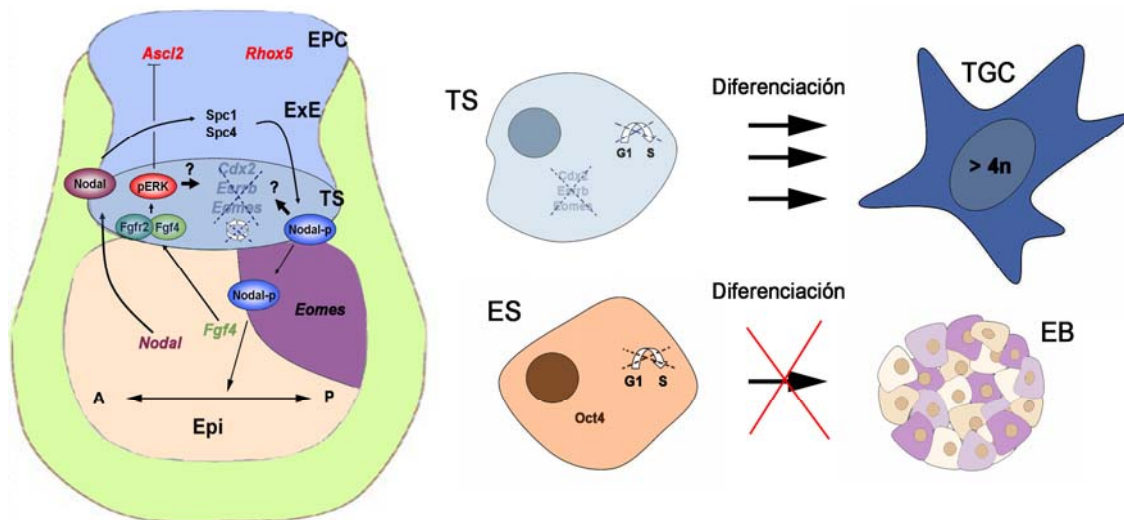


Figura 3. Fenotipo de falta de función de Dicer en el embrión temprano de ratón y células madre TS y ES. El dibujo resume el fenotipo de delección de *Dicer* descrito en este trabajo para los embriones tempranos en estadio de pos-implantación y para células TS, así como el ya descrito previamente para células ES. Todos los genes representados se encuentran expresados en los mutantes, salvo aquellos que se se que se representan tachados. En el trofoblasto los embriones mutantes para *Dicer* presentan en estadio E6.5 una pérdida específica de expresión de marcadores de células TS sin que se vea afectada la expresión génica en regiones de diferenciación (EPC). Asimismo, la proliferación se encuentra disminuida en la región del ExE donde se localizan las células TS. Este defecto no es debido a falta de señalización por parte del epiblasto, ya que tanto la vía de Fgf4 (dada por activación de Erk1/2) como de Nodal (implicada también en el establecimiento del eje antero-posterior embrionario) se encuentran activas. El epiblasto embrionario es capaz de especificar el mesodermo de forma restringida a un lado del embrión, lo que es indicativo de una correcta especificación del eje antero-posterior. En cuanto a las células madre, tanto células TS como ES tienen defectos de proliferación en ausencia de Dicer. Sin embargo, las células TS se diferencian tras la retirada de los miRNAs, mientras que las células ES son incapaces de diferenciarse cuando se induce la formación de “cuerpos embrionarios” (EB).

El análisis fenotípico de embriones mutantes para *Dicer* llevado a cabo en este trabajo muestra una correcta expresión de *Nodal*, *Fgf4* y *Eomes* en el epiblasto; asimismo resultados obtenidos en el laboratorio del Dr. T. Rodríguez muestran una correcta expresión de genes expresados en el epiblasto como *Pou5f1* (Oct4) o del gen específico de mesodermo *Brachyury* (*T*) (comunicación personal). Estos resultados demuestran cómo el embrión de ratón es capaz de establecer el patrón temprano en ausencia de miRNAs. Prueba de ello es la expresión de *Eomes* restringida a un lado del embrión mutante en estadio E7.5, que indica que el embrión no solo es capaz de especificar el mesodermo sino que además esta especificación se encuentra restringida al que probablemente será el lado posterior.

Teniendo en cuenta la importancia de los miRNAs en el mantenimiento de las poblaciones de células madre embrionarias (ES) y extraembrionarias (TS) que conforman el blastocisto resulta contradictorio el hecho de que los embriones mutantes para *Dicer* sean capaces de sobrepasar los 3.5 días de desarrollo y avanzar en el desarrollo hasta incluso especificar el mesodermo embrionario, cuando las células ES mutantes para *Dicer* son incapaces de expresar marcadores de mesodermo cuando se induce la formación de “cuerpos embrionarios” (Kanellopoulou et al., 2005). El motivo por el cual los embriones mutantes son capaces de desarrollarse más allá del estadio de blastocisto podría ser la presencia de Dicer de origen

materno durante las primeras divisiones. El requerimiento de Dicer materno en las primeras etapas del desarrollo del embrión de ratón ya se puso de manifiesto con la delección del gen en oocitos en maduración, lo que provoca una ausencia de miRNAs en el cigoto resultante en la incapacidad para progresar por a través de la primera división del embrión (Tang et al., 2007). Por este motivo parece probable que la presencia de proteína Dicer de origen materno esté contribuyendo al procesamiento de los miRNAs durante las primeras divisiones de los embriones mutantes. Un efecto similar de retraso en la aparición del fenotipo debido a la presencia de proteína Dicer de origen materno se ha descrito en pez cebra, en el que los mutantes en los que se deleciona *Dicer* tanto materno como zigótico tienen un fenotipo más temprano que el que aparece en los mutantes con delección únicamente zigótica (Giraldez et al., 2005; Wienholds et al., 2003). Los miRNAs son muy estables y no todos ellos tienen la misma vida media (Soukup et al., 2009), por lo que resulta difícil determinar en qué momento el embrión de ratón mutante para *Dicer* se encuentra libre de todos los miRNAs maduros. Así pues, familias de miRNAs altamente expresadas durante las primeras divisiones como son la familia miR17-5p o miR-291-3p (Tang et al., 2007) podrían ser procesadas por la proteína Dicer de origen materno y permanecer como miRNAs residuales en el blastocisto permitiendo que las células ES lleguen a formar el epiblasto y las TS el trofoblasto.

La importancia de la regulación por miRNAs durante el desarrollo embrionario temprano del ratón se pone también de manifiesto cuando se analizan los niveles de expresión de estos reguladores en estadios de previos a la gastrulación. En este trabajo hemos obtenido por primera vez perfiles de expresión de miRNAs durante las primeras etapas de post-implantación, lo que junto a los perfiles de expresión obtenidos para células madre ES y TS nos ha permitido localizar miRNAs con expresión predominante en los distintos linajes del embrión temprano. Muchos de estos miRNAs se encuentran formando clusters en el genoma, y parece que al menos durante estas etapas tempranas los perfiles de expresión de los integrantes de un mismo cluster son similares. Así por ejemplo, los integrantes del cluster miR-466 parecen tener expresión predominante en los linajes del trofoblasto mientras que el los miembros del cluster miR-106a analizados parecen ser específicos de linajes embrionarios.

En este trabajo hemos analizado más en profundidad el patrón de expresión varios miembros de los clusters miR-466 y miR-290, que como hemos visto en apartados anteriores parecen ser importantes para el mantenimiento de las células precursoras de los distintos linajes del embrión temprano (ES y TS). Los miembros del cluster miR-466 analizados (miR-467b y miR-669a) tienen una expresión predominante en células TS frente a ES. Cuando analizamos el patrón de expresión de estos miRNAs en el embrión temprano vemos que éste se restringe al linaje del trofoblasto y más concretamente al ectodermo extraembrionario, lo cual estaría en concordancia con una función de estos miRNAs en el mantenimiento de las células TS dentro de este tejido *in vivo* al igual que podría suceder *in vitro*.

Por otro lado el análisis por primera vez del patrón de expresión de varios miembros del cluster miR-290 en el embrión temprano muestra cómo éstos se encuentran expresados no sólo en

linajes embrionarios sino también en extraembrionarios en estadios de post-implantación. Esto concuerda con un papel específico de este cluster en el desarrollo temprano de mamíferos, ya propuesto por Houbaviv y colaboradores en 2005 (Houbaviv et al., 2005). En este trabajo se describió una regulación compleja del procesamiento de los miRNAs integrantes de este cluster, de modo que todos ellos se transcriben formando un único pri-miRNA pero luego sufren un proceso de *splicing* mediante el cual se generan distintos precursores que serán a continuación procesados por Drosha para la formación de los pre-miRNAs correspondientes (Houbaviv et al., 2005). Así, no es de extrañar que los miembros del cluster miR-290 que hemos analizado por hibridación *in situ* presenten patrones de expresión ligeramente diferentes entre sí viniendo del mismo transcrito inicial. Un patrón de expresión interesante es el que presenta miR-291-3p, que se encuentra entre los más relevantes en la regulación de las células madre del blastocisto. Este miRNA se expresa de forma homogénea en el epiblasto y trofoblasto en estadio E5.5, mientras que en E6.5 la expresión en ambos linajes persiste pero en el caso del epiblasto se encuentra restringida a un lado del embrión. La expresión ubicua en ambos linajes a los 5.5 días de desarrollo podría reflejar el papel de miR-291-3p en el mantenimiento de las poblaciones celulares más indiferenciadas durante las etapas previas al establecimiento del eje antero-posterior. Sin embargo en el embrión en estadio E6.5 se ha producido ya la migración del endodermo visceral anterior (AVE) y la consiguiente especificación del eje antero-posterior, que trae consigo el comienzo de la formación del mesodermo en la región posterior del embrión (Tam and Loebel, 2007). En este estadio sería difícil adjudicar a miR-291-3p un papel en el mantenimiento de la multipotencia, sin embargo su expresión restringida a un lado del embrión podría ser el reflejo de una función en los eventos de formación de patrón que están teniendo lugar en este momento. Entre los genes predichos como regulados por miR-291-3p se encuentra *Lefty1*, un antagonista de la vía de Nodal cuya expresión en el AVE contribuye a la restricción de Nodal a la región anterior del embrión durante la especificación del eje antero-posterior (Tam and Loebel, 2007). Los ortólogos de miR-302, miembro de la familia de miR-291-3p, se han implicado en la regulación de *Lefty1* durante el desarrollo de pez cebra y xenopus (Choi et al., 2007; Rosa et al., 2009); sin embargo miR-302 presenta expresión ubicua en el embrión de ratón a los 6.5 días de desarrollo (Card et al., 2008), mientras que miR-291-3p se encuentra restringido a un lado del embrión. Para la correcta formación del patrón en el embrión temprano la expresión de *Lefty1* ha de estar restringida al AVE en E6.5, por lo que miR-291-3p expresado en la región complementaria (posterior) podría actuar confiriendo un grado extra de represión del gen en esta zona y al mismo tiempo perfilando su territorio de expresión en la zona anterior.

6.1 miRNAs como moduladores de la expresión génica durante el desarrollo

El análisis del fenotipo del mutante de *Dicer* nos ha permitido comprobar la importancia de la función de los miRNAs durante el desarrollo embrionario temprano en ratón, sin embargo también pone de manifiesto un papel de estos reguladores durante el desarrollo en el sentido

de modular la expresión génica y contribuir aportando un grado extra de solidez a patrones de expresión que se encuentran regulados de forma principal por otros medios como son los factores de transcripción. Prueba de ello es el hecho de que no encontramos defectos en el establecimiento del patrón temprano del epiblasto, que llega a establecer el eje antero-posterior en ausencia de miRNAs. De forma similar, la delección de *Dicer* en la extremidad o el pulmón del embrión en desarrollo tampoco afecta a la formación del patrón de los mismos (Harfe et al., 2005; Harris et al., 2006). Es de destacar el hecho de que no hemos encontrado ningún gen sobre-expresado entre los analizados por hibridación *in situ*, que sería lo esperable tras la delección de un represor. Esto refleja una función general de los miRNAs durante el desarrollo temprano diferente de la que ejercen los factores de transcripción, de modo que la represión mediada por miRNAs no va encaminada a apagar la expresión génica, sino que colabora en el mantenimiento de un patrón correcto de expresión mediante la modulación precisa de la cantidad final de proteína, al mismo tiempo que colabora en la represión ejercida por otros mecanismos como son los factores de transcripción. Ejemplo de ello es el hecho de que se han generado ratones mutantes para los tres clusters que cubren a la familia miR-17-5p así como el cluster miR-290 y sólo los mutantes para el cluster miR-290 son letales embrionarios, aunque algunos de ellos llegan incluso a adultos (Ambros and Chen, 2007; Ventura et al., 2008). Estos clusters contienen miRNAs que se encuentran entre los más expresados en el embrión temprano y forman parte además de las familias más representativas durante estos estadios, sin embargo no parece que sean imprescindibles para una correcta especificación del patrón temprano, lo que apoya la idea de los miRNAs como moduladores de la expresión génica más que como mecanismos determinantes de ella en los procesos de desarrollo. Estudios enfocados al análisis de la función de miRNAs concretos serán necesarios para conocer en profundidad en qué medida la regulación mediada por estos represores contribuye a los distintos eventos que tienen lugar durante las primeras etapas del desarrollo embrionario en ratón.

Conclusiones

1. El mantenimiento de la expresión de los genes *Cdx2*, *Esrrb* y *Eomes*, responsables de multipotencia de la población de células madre del trofoblasto (células TS) en el ectodermo extraembrionario depende de miRNAs.
2. Los miRNAs se requieren para el mantenimiento de la proliferación en el ectodermo extraembrionario, donde se localiza la población de células TS.
3. El establecimiento del patrón embrionario temprano en el epiblasto es independiente de miRNAs.
4. La delección de *Dicer* en células TS en cultivo trae consigo una bajada drástica de la proliferación, acompañada de la diferenciación en los distintos tipos celulares del linaje del trofoblasto.
5. El fenotipo observado en células TS mutantes para *Dicer* en cultivo refleja los resultados obtenidos del análisis en el embrión temprano, y representa así un buen modelo para el estudio de la regulación mediada por miRNAs en el linaje del trofoblasto.
6. La caracterización de la expresión de miRNAs en células TS y embriones en estadios tempranos de post-implantación ha servido para identificar familias de miRNAs expresados específicamente en el linaje del trofoblasto, así como familias específicas de embrión temprano y células madre derivadas del blastocisto (ES y TS).
7. Los miRNAs colaboran en la proliferación de las células TS mediante la represión de los inhibidores de la transición G1 a S del ciclo celular p21, Rbl2 y Lats2 de forma similar a como lo hacen en células ES. Las familias de miRNAs implicadas en esta represión son comunes a otros tipos celulares (miR-17-5p) o específicas de células madre derivadas del blastocisto (miR-291-3p).
8. Los miRNAs contribuyen a la multipotencia de las células TS en cultivo mediante la represión de factores implicados específicamente en la diferenciación del linaje del trofoblasto. La regulación por distintas familias de miRNAs, entre las que se encuentran algunas específicas del linaje del trofoblasto como miR-466 o miR-467b, podría contribuir a la represión de la diferenciación hacia células gigantes del trofoblasto mediante el control de la endo-reduplicación mediada por p57.
9. El mecanismo de actuación que se infiere del análisis del fenotipo de embriones y células madre derivadas del blastocisto mutantes de *Dicer* indica que la función de los miRNAs está encaminada a la modulación fina de la expresión génica, contribuyendo al mantenimiento de patrones de expresión establecidos fundamentalmente a nivel transcripcional.

Bibliografía

- Abramoff, M.D., Magelhaes, P.J., and Ram, S.J. (2004). Image Processing with ImageJ. *Biophotonics International* 11, 36-42.
- Al-Shahrour, F., Minguez, P., Tarraga, J., Montaner, D., Alloza, E., Vaquerizas, J.M., Conde, L., Blaschke, C., Vera, J., and Dopazo, J. (2006). BABELOMICS: a systems biology perspective in the functional annotation of genome-scale experiments. *Nucleic Acids Res* 34, W472-476.
- Al-Shahrour, F., Minguez, P., Vaquerizas, J.M., Conde, L., and Dopazo, J. (2005). BABELOMICS: a suite of web tools for functional annotation and analysis of groups of genes in high-throughput experiments. *Nucleic Acids Res* 33, W460-464.
- Ambros, V. (2004). The functions of animal microRNAs. *Nature* 431, 350-355.
- Ambros, V., and Chen, X. (2007). The regulation of genes and genomes by small RNAs. *Development* 134, 1635-1641.
- Ariza-McNaughton, L., and Krumlauf, R. (2002). Non-radioactive in situ hybridization: simplified procedures for use in whole-mounts of mouse and chick embryos. *Int Rev Neurobiol* 47, 239-250.
- Arman, E., Haffner-Krausz, R., Chen, Y., Heath, J.K., and Lonai, P. (1998). Targeted disruption of fibroblast growth factor (FGF) receptor 2 suggests a role for FGF signaling in pregastrulation mammalian development. *Proc Natl Acad Sci U S A* 95, 5082-5087.
- Arnold, S.J., Hofmann, U.K., Bikoff, E.K., and Robertson, E.J. (2008). Pivotal roles for eomesodermin during axis formation, epithelium-to-mesenchyme transition and endoderm specification in the mouse. *Development* 135, 501-511.
- Babiarz, J.E., Ruby, J.G., Wang, Y., Bartel, D.P., and Blelloch, R. (2008). Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev* 22, 2773-2785.
- Baek, D., Villen, J., Shin, C., Camargo, F.D., Gygi, S.P., and Bartel, D.P. (2008). The impact of microRNAs on protein output. *Nature* 455, 64-71.
- Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281-297.
- Baskerville, S., and Bartel, D.P. (2005). Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA* 11, 241-247.
- Benetti, R., Gonzalo, S., Jaco, I., Munoz, P., Gonzalez, S., Schoeftner, S., Murchison, E., Andl, T., Chen, T., Klatt, P., *et al.* (2008). A mammalian microRNA cluster controls DNA methylation and telomere recombination via Rbl2-dependent regulation of DNA methyltransferases. *Nat Struct Mol Biol* 15, 268-279.
- Berezikov, E., Chung, W.J., Willis, J., Cuppen, E., and Lai, E.C. (2007). Mammalian mirtron genes. *Mol Cell* 28, 328-336.
- Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409, 363-366.
- Bernstein, E., Kim, S.Y., Carmell, M.A., Murchison, E.P., Alcorn, H., Li, M.Z., Mills, A.A., Elledge, S.J., Anderson, K.V., and Hannon, G.J. (2003). Dicer is essential for mouse development. *Nat Genet* 35, 215-217.
- Betel, D., Wilson, M., Gabow, A., Marks, D.S., and Sander, C. (2008). The microRNA.org resource: targets and expression. *Nucleic Acids Res* 36, D149-153.
- Blagosklonny, M.V., and Pardee, A.B. (2002). The restriction point of the cell cycle. *Cell Cycle* 1, 103-110.

Borchert, G.M., Lanier, W., and Davidson, B.L. (2006). RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol* 13, 1097-1101.

Brennecke, J., Hipfner, D.R., Stark, A., Russell, R.B., and Cohen, S.M. (2003). bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in *Drosophila*. *Cell* 113, 25-36.

Bueno, M.J., Perez de Castro, I., Gomez de Cedron, M., Santos, J., Calin, G.A., Cigudosa, J.C., Croce, C.M., Fernandez-Piqueras, J., and Malumbres, M. (2008). Genetic and epigenetic silencing of microRNA-203 enhances ABL1 and BCR-ABL1 oncogene expression. *Cancer Cell* 13, 496-506.

Calabrese, J.M., Seila, A.C., Yeo, G.W., and Sharp, P.A. (2007). RNA sequence analysis defines Dicer's role in mouse embryonic stem cells. *Proc Natl Acad Sci U S A* 104, 18097-18102.

Card, D.A., Hebbar, P.B., Li, L., Trotter, K.W., Komatsu, Y., Mishina, Y., and Archer, T.K. (2008). Oct4/Sox2-regulated miR-302 targets cyclin D1 in human embryonic stem cells. *Mol Cell Biol* 28, 6426-6438.

Ciruna, B.G., and Rossant, J. (1999). Expression of the T-box gene Eomesodermin during early mouse development. *Mech Dev* 81, 199-203.

Cobb, B.S., Nesterova, T.B., Thompson, E., Hertweck, A., O'Connor, E., Godwin, J., Wilson, C.B., Brockdorff, N., Fisher, A.G., Smale, S.T., *et al.* (2005). T cell lineage choice and differentiation in the absence of the RNase III enzyme Dicer. *J Exp Med* 201, 1367-1373.

Conlon, F.L., Lyons, K.M., Takaesu, N., Barth, K.S., Kispert, A., Herrmann, B., and Robertson, E.J. (1994). A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. *Development* 120, 1919-1928.

Cross, J.C. (2000). Genetic insights into trophoblast differentiation and placental morphogenesis. *Semin Cell Dev Biol* 11, 105-113.

Chen, K., and Rajewsky, N. (2007). The evolution of gene regulation by transcription factors and microRNAs. *Nat Rev Genet* 8, 93-103.

Chendrimada, T.P., Gregory, R.I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., and Shiekhattar, R. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* 436, 740-744.

Choi, W.Y., Giraldez, A.J., and Schier, A.F. (2007). Target protectors reveal dampening and balancing of Nodal agonist and antagonist by miR-430. *Science* 318, 271-274.

Christie, G.R., Williams, D.J., Macisaac, F., Dickinson, R.J., Rosewell, I., and Keyse, S.M. (2005). The dual-specificity protein phosphatase DUSP9/MKP-4 is essential for placental function but is not required for normal embryonic development. *Mol Cell Biol* 25, 8323-8333.

Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. *Nature* 432, 231-235.

Erlebacher, A., Price, K.A., and Glimcher, L.H. (2004). Maintenance of mouse trophoblast stem cell proliferation by TGF-beta/activin. *Dev Biol* 275, 158-169.

Filipowicz, W., Bhattacharyya, S.N., and Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet* 9, 102-114.

Foshay, K.M., and Gallicano, G.I. (2009). miR-17 family miRNAs are expressed during early mammalian development and regulate stem cell differentiation. *Dev Biol* 326, 431-443.

Frankenberg, S., Smith, L., Greenfield, A., and Zernicka-Goetz, M. (2007). Novel gene expression patterns along the proximo-distal axis of the mouse embryo before gastrulation. *BMC Dev Biol* 7, 8.

Ghildiyal, M., and Zamore, P.D. (2009). Small silencing RNAs: an expanding universe. *Nat Rev Genet* 10, 94-108.

Giraldez, A.J., Cinalli, R.M., Glasner, M.E., Enright, A.J., Thomson, J.M., Baskerville, S., Hammond, S.M., Bartel, D.P., and Schier, A.F. (2005). MicroRNAs regulate brain morphogenesis in zebrafish. *Science* 308, 833-838.

Giraldez, A.J., Mishima, Y., Rihel, J., Grocock, R.J., Van Dongen, S., Inoue, K., Enright, A.J., and Schier, A.F. (2006). Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* 312, 75-79.

Griffiths-Jones, S., Saini, H.K., van Dongen, S., and Enright, A.J. (2008). miRBase: tools for microRNA genomics. *Nucleic Acids Res* 36, D154-158.

Grimson, A., Farh, K.K., Johnston, W.K., Garrett-Engele, P., Lim, L.P., and Bartel, D.P. (2007). MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol Cell* 27, 91-105.

Guillemot, F., Nagy, A., Auerbach, A., Rossant, J., and Joyner, A.L. (1994). Essential role of Mash-2 in extraembryonic development. *Nature* 371, 333-336.

Guzman-Ayala, M., Ben-Haim, N., Beck, S., and Constam, D.B. (2004). Nodal protein processing and fibroblast growth factor 4 synergize to maintain a trophoblast stem cell microenvironment. *Proc Natl Acad Sci U S A* 101, 15656-15660.

Harfe, B.D., McManus, M.T., Mansfield, J.H., Hornstein, E., and Tabin, C.J. (2005). The RNaseIII enzyme Dicer is required for morphogenesis but not patterning of the vertebrate limb. *Proc Natl Acad Sci U S A* 102, 10898-10903.

Harris, K.S., Zhang, Z., McManus, M.T., Harfe, B.D., and Sun, X. (2006). Dicer function is essential for lung epithelium morphogenesis. *Proc Natl Acad Sci U S A* 103, 2208-2213.

Harvey, M.B., Leco, K.J., Arcellana-Panlilio, M.Y., Zhang, X., Edwards, D.R., and Schultz, G.A. (1995). Proteinase expression in early mouse embryos is regulated by leukaemia inhibitory factor and epidermal growth factor. *Development* 121, 1005-1014.

Hatfield, S.D., Shcherbata, H.R., Fischer, K.A., Nakahara, K., Carthew, R.W., and Ruohola-Baker, H. (2005). Stem cell division is regulated by the microRNA pathway. *Nature* 435, 974-978.

Hayashi, S., Lewis, P., Pevny, L., and McMahon, A.P. (2002). Efficient gene modulation in mouse epiblast using a Sox2Cre transgenic mouse strain. *Gene Expr Patterns* 2, 93-97.

He, L., He, X., Lim, L.P., de Stanchina, E., Xuan, Z., Liang, Y., Xue, W., Zender, L., Magnus, J., Ridzon, D., *et al.* (2007). A microRNA component of the p53 tumour suppressor network. *Nature* 447, 1130-1134.

He, L., Thomson, J.M., Hemann, M.T., Hernando-Monge, E., Mu, D., Goodson, S., Powers, S., Cordon-Cardo, C., Lowe, S.W., Hannon, G.J., *et al.* (2005). A microRNA polycistron as a potential human oncogene. *Nature* 435, 828-833.

Hemberger, M. (2007). Epigenetic landscape required for placental development. *Cell Mol Life Sci* 64, 2422-2436.

Hemberger, M., and Cross, J.C. (2001). Genes governing placental development. *Trends Endocrinol Metab* 12, 162-168.

- Himeno, E., Tanaka, S., and Kunath, T. (2008). Isolation and manipulation of mouse trophoblast stem cells. *Curr Protoc Stem Cell Biol Chapter 1*, Unit 1E 4.
- Hornstein, E., and Shomron, N. (2006). Canalization of development by microRNAs. *Nat Genet 38 Suppl*, S20-24.
- Houbaviy, H.B., Dennis, L., Jaenisch, R., and Sharp, P.A. (2005). Characterization of a highly variable eutherian microRNA gene. *RNA 11*, 1245-1257.
- Houbaviy, H.B., Murray, M.F., and Sharp, P.A. (2003). Embryonic stem cell-specific MicroRNAs. *Dev Cell 5*, 351-358.
- Hughes, M., Dobric, N., Scott, I.C., Su, L., Starovic, M., St-Pierre, B., Egan, S.E., Kingdom, J.C., and Cross, J.C. (2004). The Hand1, Stra13 and Gcm1 transcription factors override FGF signaling to promote terminal differentiation of trophoblast stem cells. *Dev Biol 271*, 26-37.
- Hwang, H.W., Wentzel, E.A., and Mendell, J.T. (2007). A hexanucleotide element directs microRNA nuclear import. *Science 315*, 97-100.
- Ivanovska, I., Ball, A.S., Diaz, R.L., Magnus, J.F., Kibukawa, M., Schelter, J.M., Kobayashi, S.V., Lim, L., Burchard, J., Jackson, A.L., *et al.* (2008). MicroRNAs in the miR-106b family regulate p21/CDKN1A and promote cell cycle progression. *Mol Cell Biol 28*, 2167-2174.
- Jacquemin, P., Sapin, V., Alsat, E., Evain-Brion, D., Dolle, P., and Davidson, I. (1998). Differential expression of the TEF family of transcription factors in the murine placenta and during differentiation of primary human trophoblasts in vitro. *Dev Dyn 212*, 423-436.
- Kanellopoulou, C., Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M., and Rajewsky, K. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev 19*, 489-501.
- Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev 15*, 2654-2659.
- Khvorova, A., Reynolds, A., and Jayasena, S.D. (2003). Functional siRNAs and miRNAs exhibit strand bias. *Cell 115*, 209-216.
- Kloosterman, W.P., and Plasterk, R.H. (2006). The diverse functions of microRNAs in animal development and disease. *Dev Cell 11*, 441-450.
- Kloosterman, W.P., Wienholds, E., de Bruijn, E., Kauppinen, S., and Plasterk, R.H. (2006). In situ detection of miRNAs in animal embryos using LNA-modified oligonucleotide probes. *Nat Methods 3*, 27-29.
- Krek, A., Grun, D., Poy, M.N., Wolf, R., Rosenberg, L., Epstein, E.J., MacMenamin, P., da Piedade, I., Gunsalus, K.C., Stoffel, M., *et al.* (2005). Combinatorial microRNA target predictions. *Nat Genet 37*, 495-500.
- Kunath, T., Strumpf, D., and Rossant, J. (2004). Early trophoblast determination and stem cell maintenance in the mouse--a review. *Placenta 25 Suppl A*, S32-38.
- Lagos-Quintana, M., Rauhut, R., Meyer, J., Borkhardt, A., and Tuschl, T. (2003). New microRNAs from mouse and human. *RNA 9*, 175-179.
- Lagos-Quintana, M., Rauhut, R., Yalcin, A., Meyer, J., Lendeckel, W., and Tuschl, T. (2002). Identification of tissue-specific microRNAs from mouse. *Curr Biol 12*, 735-739.

- Landgraf, P., Rusu, M., Sheridan, R., Sewer, A., Iovino, N., Aravin, A., Pfeffer, S., Rice, A., Kamphorst, A.O., Landthaler, M., *et al.* (2007). A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* 129, 1401-1414.
- Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75, 843-854.
- Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H., and Kim, V.N. (2004). MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* 23, 4051-4060.
- Legendre, M., Ritchie, W., Lopez, F., and Gautheret, D. (2006). Differential repression of alternative transcripts: a screen for miRNA targets. *PLoS Comput Biol* 2, e43.
- Li, M., Yee, D., Magnuson, T.R., Smithies, O., and Caron, K.M. (2006). Reduced maternal expression of adrenomedullin disrupts fertility, placentation, and fetal growth in mice. *J Clin Invest* 116, 2653-2662.
- Li, Y., Pei, J., Xia, H., Ke, H., Wang, H., and Tao, W. (2003). *Lats2*, a putative tumor suppressor, inhibits G1/S transition. *Oncogene* 22, 4398-4405.
- Lim, L.P., Lau, N.C., Garrett-Engle, P., Grimson, A., Schelter, J.M., Castle, J., Bartel, D.P., Linsley, P.S., and Johnson, J.M. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* 433, 769-773.
- Lim, L.P., Lau, N.C., Weinstein, E.G., Abdelhakim, A., Yekta, S., Rhoades, M.W., Burge, C.B., and Bartel, D.P. (2003). The microRNAs of *Caenorhabditis elegans*. *Genes Dev* 17, 991-1008.
- Lin, T.P., Labosky, P.A., Grabel, L.B., Kozak, C.A., Pitman, J.L., Kleeman, J., and MacLeod, C.L. (1994). The *Pem* homeobox gene is X-linked and exclusively expressed in extraembryonic tissues during early murine development. *Dev Biol* 166, 170-179.
- Linsley, P.S., Schelter, J., Burchard, J., Kibukawa, M., Martin, M.M., Bartz, S.R., Johnson, J.M., Cummins, J.M., Raymond, C.K., Dai, H., *et al.* (2007). Transcripts targeted by the microRNA-16 family cooperatively regulate cell cycle progression. *Mol Cell Biol* 27, 2240-2252.
- Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. (2004). *Argonaute2* is the catalytic engine of mammalian RNAi. *Science* 305, 1437-1441.
- Lu, C.W., Yabuuchi, A., Chen, L., Viswanathan, S., Kim, K., and Daley, G.Q. (2008). Ras-MAPK signaling promotes trophectoderm formation from embryonic stem cells and mouse embryos. *Nat Genet* 40, 921-926.
- Lu, J., Qian, J., Chen, F., Tang, X., Li, C., and Cardoso, W.V. (2005). Differential expression of components of the microRNA machinery during mouse organogenesis. *Biochem Biophys Res Commun* 334, 319-323.
- Lu, Y., Thomson, J.M., Wong, H.Y., Hammond, S.M., and Hogan, B.L. (2007). Transgenic over-expression of the microRNA miR-17-92 cluster promotes proliferation and inhibits differentiation of lung epithelial progenitor cells. *Dev Biol* 310, 442-453.
- Lujambio, A., and Esteller, M. (2009). How epigenetics can explain human metastasis: a new role for microRNAs. *Cell Cycle* 8, 377-382.
- Luo, J., Sladek, R., Bader, J.A., Matthysen, A., Rossant, J., and Giguere, V. (1997). Placental abnormalities in mouse embryos lacking the orphan nuclear receptor ERR-beta. *Nature* 388, 778-782.

Ma, G.T., Roth, M.E., Groskopf, J.C., Tsai, F.Y., Orkin, S.H., Grosveld, F., Engel, J.D., and Linzer, D.I. (1997). GATA-2 and GATA-3 regulate trophoblast-specific gene expression in vivo. *Development* **124**, 907-914.

Mak, W., Baxter, J., Silva, J., Newall, A.E., Otte, A.P., and Brockdorff, N. (2002). Mitotically stable association of polycomb group proteins *eed* and *enx1* with the inactive x chromosome in trophoblast stem cells. *Curr Biol* **12**, 1016-1020.

Malumbres, M., and Barbacid, M. (2009). Cell cycle, CDKs and cancer: a changing paradigm. *Nat Rev Cancer* **9**, 153-166.

Mansfield, J.H., Harfe, B.D., Nissen, R., Obenauer, J., Srineel, J., Chaudhuri, A., Farzan-Kashani, R., Zuker, M., Pasquinelli, A.E., Ruvkun, G., *et al.* (2004). MicroRNA-responsive 'sensor' transgenes uncover Hox-like and other developmentally regulated patterns of vertebrate microRNA expression. *Nat Genet* **36**, 1079-1083.

Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J., *et al.* (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* **134**, 521-533.

Martinez Barbera, J.P., Clements, M., Thomas, P., Rodriguez, T., Meloy, D., Kioussis, D., and Beddington, R.S. (2000). The homeobox gene *Hex* is required in definitive endodermal tissues for normal forebrain, liver and thyroid formation. *Development* **127**, 2433-2445.

Maziere, P., and Enright, A.J. (2007). Prediction of microRNA targets. *Drug Discov Today* **12**, 452-458.

Melo, S.A., Ropero, S., Moutinho, C., Aaltonen, L.A., Yamamoto, H., Calin, G.A., Rossi, S., Fernandez, A.F., Carneiro, F., Oliveira, C., *et al.* (2009). A TARBP2 mutation in human cancer impairs microRNA processing and DICER1 function. *Nat Genet* **41**, 365-370.

Mendell, J.T. (2008). miRiad roles for the miR-17-92 cluster in development and disease. *Cell* **133**, 217-222.

Miska, E.A., Alvarez-Saavedra, E., Abbott, A.L., Lau, N.C., Hellman, A.B., McGonagle, S.M., Bartel, D.P., Ambros, V.R., and Horvitz, H.R. (2007). Most *Caenorhabditis elegans* microRNAs are individually not essential for development or viability. *PLoS Genet* **3**, e215.

Molnar, A., Schwach, F., Studholme, D.J., Thuenemann, E.C., and Baulcombe, D.C. (2007). miRNAs control gene expression in the single-cell alga *Chlamydomonas reinhardtii*. *Nature* **447**, 1126-1129.

Morita, S., Horii, T., Kimura, M., Goto, Y., Ochiya, T., and Hatada, I. (2007). One Argonaute family member, *Eif2c2* (*Ago2*), is essential for development and appears not to be involved in DNA methylation. *Genomics* **89**, 687-696.

Morton, S.U., Scherz, P.J., Cordes, K.R., Ivey, K.N., Stainier, D.Y., and Srivastava, D. (2008). microRNA-138 modulates cardiac patterning during embryonic development. *Proc Natl Acad Sci U S A* **105**, 17830-17835.

Mudhasani, R., Zhu, Z., Hutvagner, G., Eischen, C.M., Lyle, S., Hall, L.L., Lawrence, J.B., Imbalzano, A.N., and Jones, S.N. (2008). Loss of miRNA biogenesis induces p19Arf-p53 signaling and senescence in primary cells. *J Cell Biol* **181**, 1055-1063.

Murchison, E.P., Partridge, J.F., Tam, O.H., Cheloufi, S., and Hannon, G.J. (2005). Characterization of Dicer-deficient murine embryonic stem cells. *Proc Natl Acad Sci U S A* **102**, 12135-12140.

Murchison, E.P., Stein, P., Xuan, Z., Pan, H., Zhang, M.Q., Schultz, R.M., and Hannon, G.J. (2007). Critical roles for Dicer in the female germline. *Genes Dev* **21**, 682-693.

- Nagy, A., Gertsenstein, M., Vintersen, K., and Behringer, R. (2003). Manipulating the mouse embryo. A laboratory manual, 3rd edn (Cold Spring Harbor, New York, Cold Spring Harbor Laboratory Press).
- Nakayama, H., Scott, I.C., and Cross, J.C. (1998). The transition to endoreduplication in trophoblast giant cells is regulated by the mSNA zinc finger transcription factor. *Dev Biol* **199**, 150-163.
- Ng, R.K., Dean, W., Dawson, C., Lucifero, D., Madeja, Z., Reik, W., and Hemberger, M. (2008). Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5. *Nat Cell Biol* **10**, 1280-1290.
- Nishioka, N., Inoue, K., Adachi, K., Kiyonari, H., Ota, M., Ralston, A., Yabuta, N., Hirahara, S., Stephenson, R.O., Ogonuki, N., *et al.* (2009). The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. *Dev Cell* **16**, 398-410.
- Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., and Rossant, J. (2005). Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* **123**, 917-929.
- Nose, A., and Takeichi, M. (1986). A novel cadherin cell adhesion molecule: its expression patterns associated with implantation and organogenesis of mouse embryos. *J Cell Biol* **103**, 2649-2658.
- O'Donnell, K.A., Wentzel, E.A., Zeller, K.I., Dang, C.V., and Mendell, J.T. (2005). c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **435**, 839-843.
- O'Rourke, J.R., Georges, S.A., Seay, H.R., Tapscott, S.J., McManus, M.T., Goldhamer, D.J., Swanson, M.S., and Harfe, B.D. (2007). Essential role for Dicer during skeletal muscle development. *Dev Biol* **311**, 359-368.
- Pasquinelli, A.E., McCoy, A., Jimenez, E., Salo, E., Ruvkun, G., Martindale, M.Q., and Baguna, J. (2003). Expression of the 22 nucleotide let-7 heterochronic RNA throughout the Metazoa: a role in life history evolution? *Evol Dev* **5**, 372-378.
- Pasquinelli, A.E., Reinhart, B.J., Slack, F., Martindale, M.Q., Kuroda, M.I., Maller, B., Hayward, D.C., Ball, E.E., Degan, B., Muller, P., *et al.* (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* **408**, 86-89.
- Patterson, K.I., Brummer, T., O'Brien, P.M., and Daly, R.J. (2009). Dual-specificity phosphatases: critical regulators with diverse cellular targets. *Biochem J* **418**, 475-489.
- Pierson, J., Hostager, B., Fan, R., and Vibhakkar, R. (2008). Regulation of cyclin dependent kinase 6 by microRNA 124 in medulloblastoma. *J Neurooncol* **90**, 1-7.
- Sun, F., Fu, H., Liu, Q., Tie, Y., Zhu, J., Xing, R., Sun, Z., and Zheng, X. (2008). Downregulation of CCND1 and CDK6 by miR-34a induces cell cycle arrest. *FEBS Lett* **582**, 1564-1568.
- Plasterk, R.H. (2006). Micro RNAs in animal development. *Cell* **124**, 877-881.
- Ravid, K., Lu, J., Zimet, J.M., and Jones, M.R. (2002). Roads to polyploidy: the megakaryocyte example. *J Cell Physiol* **190**, 7-20.
- Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R., and Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* **403**, 901-906.
- Rielland, M., Hue, I., Renard, J.P., and Alice, J. (2008). Trophoblast stem cell derivation, cross-species comparison and use of nuclear transfer: new tools to study trophoblast growth and differentiation. *Dev Biol* **322**, 1-10.

Rodriguez, A., Griffiths-Jones, S., Ashurst, J.L., and Bradley, A. (2004). Identification of mammalian microRNA host genes and transcription units. *Genome Res* 14, 1902-1910.

Rodriguez, T.A., Srinivas, S., Clements, M.P., Smith, J.C., and Beddington, R.S. (2005). Induction and migration of the anterior visceral endoderm is regulated by the extra-embryonic ectoderm. *Development* 132, 2513-2520.

Roebroek, A.J., Umans, L., Pauli, I.G., Robertson, E.J., van Leuven, F., Van de Ven, W.J., and Constam, D.B. (1998). Failure of ventral closure and axial rotation in embryos lacking the proprotein convertase Furin. *Development* 125, 4863-4876.

Roper, S., and Hemberger, M. (2009). Defining pathways that enforce cell lineage specification in early development and stem cells. *Cell Cycle* 8.

Rosa, A., Spagnoli, F.M., and Brivanlou, A.H. (2009). The miR-430/427/302 family controls mesendodermal fate specification via species-specific target selection. *Dev Cell* 16, 517-527.

Rossant, J. (2004). Lineage development and polar asymmetries in the peri-implantation mouse blastocyst. *Semin Cell Dev Biol* 15, 573-581.

Rossant, J., and Cross, J.C. (2001). Placental development: lessons from mouse mutants. *Nat Rev Genet* 2, 538-548.

Rossant, J., and Tam, P.P. (2009). Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. *Development* 136, 701-713.

Ruby, J.G., Jan, C.H., and Bartel, D.P. (2007). Intronic microRNA precursors that bypass Drosha processing. *Nature* 448, 83-86.

Russ, A.P., Wattler, S., Colledge, W.H., Aparicio, S.A., Carlton, M.B., Pearce, J.J., Barton, S.C., Surani, M.A., Ryan, K., Nehls, M.C., *et al.* (2000). Eomesodermin is required for mouse trophoblast development and mesoderm formation. *Nature* 404, 95-99.

Sandberg, R., Neilson, J.R., Sarma, A., Sharp, P.A., and Burge, C.B. (2008). Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science* 320, 1643-1647.

Sawai, S., Shimono, A., Hanaoka, K., and Kondoh, H. (1991). Embryonic lethality resulting from disruption of both N-myc alleles in mouse zygotes. *New Biol* 3, 861-869.

Schenke-Layland, K., Angelis, E., Rhodes, K.E., Heydarkhan-Hagvall, S., Mikkola, H.K., and Maclellan, W.R. (2007). Collagen IV induces trophoectoderm differentiation of mouse embryonic stem cells. *Stem Cells* 25, 1529-1538.

Schultz, R.M. (2002). The molecular foundations of the maternal to zygotic transition in the preimplantation embryo. *Hum Reprod Update* 8, 323-331.

Seitz, H., Royo, H., Bortolin, M.L., Lin, S.P., Ferguson-Smith, A.C., and Cavaille, J. (2004). A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. *Genome Res* 14, 1741-1748.

Selbach, M., Schwanhaussner, B., Thierfelder, N., Fang, Z., Khanin, R., and Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. *Nature* 455, 58-63.

Shkumatava, A., Stark, A., Sive, H., and Bartel, D.P. (2009). Coherent but overlapping expression of microRNAs and their targets during vertebrate development. *Genes Dev* 23, 466-481.

Sinkkonen, L., Hugenschmidt, T., Berninger, P., Gaidatzis, D., Mohn, F., Artus-Revel, C.G., Zavolan, M., Svoboda, P., and Filipowicz, W. (2008). MicroRNAs control de novo DNA

methylation through regulation of transcriptional repressors in mouse embryonic stem cells. *Nat Struct Mol Biol* 15, 259-267.

Song, G., and Wang, L. (2008). MiR-433 and miR-127 arise from independent overlapping primary transcripts encoded by the miR-433-127 locus. *PLoS ONE* 3, e3574.

Soukup, G.A., Frittsch, B., Pierce, M.L., Weston, M.D., Jahan, I., McManus, M.T., and Harfe, B.D. (2009). Residual microRNA expression dictates the extent of inner ear development in conditional Dicer knockout mice. *Dev Biol* 328, 328-341.

Stark, A., Brennecke, J., Bushati, N., Russell, R.B., and Cohen, S.M. (2005). Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. *Cell* 123, 1133-1146.

Stefani, G., and Slack, F.J. (2008). Small non-coding RNAs in animal development. *Nat Rev Mol Cell Biol* 9, 219-230.

Strumpf, D., Mao, C.A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., and Rossant, J. (2005). *Cdx2* is required for correct cell fate specification and differentiation of trophoblast in the mouse blastocyst. *Development* 132, 2093-2102.

Takahashi, Y., Carpino, N., Cross, J.C., Torres, M., Parganas, E., and Ihle, J.N. (2003). SOCS3: an essential regulator of LIF receptor signaling in trophoblast giant cell differentiation. *EMBO J* 22, 372-384.

Takahashi, Y., Takahashi, M., Carpino, N., Jou, S.T., Chao, J.R., Tanaka, S., Shigeyoshi, Y., Parganas, E., and Ihle, J.N. (2008). Leukemia inhibitory factor regulates trophoblast giant cell differentiation via Janus kinase 1-signal transducer and activator of transcription 3-suppressor of cytokine signaling 3 pathway. *Mol Endocrinol* 22, 1673-1681.

Tam, P.P., and Loebel, D.A. (2007). Gene function in mouse embryogenesis: get set for gastrulation. *Nat Rev Genet* 8, 368-381.

Tanaka, S., Kunath, T., Hadjantonakis, A.K., Nagy, A., and Rossant, J. (1998). Promotion of trophoblast stem cell proliferation by FGF4. *Science* 282, 2072-2075.

Tang, F., Kaneda, M., O'Carroll, D., Hajkova, P., Barton, S.C., Sun, Y.A., Lee, C., Tarakhovsky, A., Lao, K., and Surani, M.A. (2007). Maternal microRNAs are essential for mouse zygotic development. *Genes Dev* 21, 644-648.

Tay, Y., Zhang, J., Thomson, A.M., Lim, B., and Rigoutsos, I. (2008). MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature* 455, 1124-1128.

Ullah, Z., Kohn, M.J., Yagi, R., Vassilev, L.T., and DePamphilis, M.L. (2008). Differentiation of trophoblast stem cells into giant cells is triggered by p57/Kip2 inhibition of CDK1 activity. *Genes Dev* 22, 3024-3036.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3, RESEARCH0034.

Vasudevan, S., Tong, Y., and Steitz, J.A. (2007). Switching from repression to activation: microRNAs can up-regulate translation. *Science* 318, 1931-1934.

Ventura, A., and Jacks, T. (2009). MicroRNAs and cancer: short RNAs go a long way. *Cell* 136, 586-591.

Ventura, A., Young, A.G., Winslow, M.M., Lintault, L., Meissner, A., Erkland, S.J., Newman, J., Bronson, R.T., Crowley, D., Stone, J.R., *et al.* (2008). Targeted deletion reveals essential and overlapping functions of the miR-17 through 92 family of miRNA clusters. *Cell* 132, 875-886.

Wang, Y., Baskerville, S., Shenoy, A., Babiarz, J.E., Baehner, L., and Bluelloch, R. (2008). Embryonic stem cell-specific microRNAs regulate the G1-S transition and promote rapid proliferation. *Nat Genet* 40, 1478-1483.

Wang, Y., Medvid, R., Melton, C., Jaenisch, R., and Bluelloch, R. (2007). DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. *Nat Genet* 39, 380-385.

Wienholds, E., Koudijs, M.J., van Eeden, F.J., Cuppen, E., and Plasterk, R.H. (2003). The microRNA-producing enzyme Dicer1 is essential for zebrafish development. *Nat Genet* 35, 217-218.

Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* 75, 855-862.

Wilkinson, D. (1992). Whole mount in situ hybridisation of vertebrate embryos.

Winter, J., Jung, S., Keller, S., Gregory, R.I., and Diederichs, S. (2009). Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat Cell Biol* 11, 228-234.

Yagi, R., Kohn, M.J., Karavanova, I., Kaneko, K.J., Vullhorst, D., DePamphilis, M.L., and Buonanno, A. (2007). Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development. *Development* 134, 3827-3836.

Yang, W., Klamman, L.D., Chen, B., Araki, T., Harada, H., Thomas, S.M., George, E.L., and Neel, B.G. (2006). An Shp2/SFK/Ras/Erk signaling pathway controls trophoblast stem cell survival. *Dev Cell* 10, 317-327.

Yang, W.J., Yang, D.D., Na, S., Sandusky, G.E., Zhang, Q., and Zhao, G. (2005). Dicer is required for embryonic angiogenesis during mouse development. *J Biol Chem* 280, 9330-9335.

Yarwood, S., Bouyoucef-Cherchalli, D., Cullen, P.J., and Kupzig, S. (2006). The GAP1 family of GTPase-activating proteins: spatial and temporal regulators of small GTPase signalling. *Biochem Soc Trans* 34, 846-850.

Yi, R., Pasolli, H.A., Landthaler, M., Hafner, M., Ojo, T., Sheridan, R., Sander, C., O'Carroll, D., Stoffel, M., Tuschl, T., *et al.* (2009). DGCR8-dependent microRNA biogenesis is essential for skin development. *Proc Natl Acad Sci U S A* 106, 498-502.

Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* 17, 3011-3016.

Yoo, A.S., and Greenwald, I. (2005). LIN-12/Notch activation leads to microRNA-mediated down-regulation of *Vav* in *C. elegans*. *Science* 310, 1330-1333.

Yu, J.Y., Reynolds, S.H., Hatfield, S.D., Shcherbata, H.R., Fischer, K.A., Ward, E.J., Long, D., Ding, Y., and Ruohola-Baker, H. (2009). Dicer-1-dependent Dacapo suppression acts downstream of Insulin receptor in regulating cell division of *Drosophila* germline stem cells. *Development* 136, 1497-1507.

Zhang, B., Pan, X., Cobb, G.P., and Anderson, T.A. (2007). microRNAs as oncogenes and tumor suppressors. *Dev Biol* 302, 1-12.

Zhao, T., Li, G., Mi, S., Li, S., Hannon, G.J., Wang, X.J., and Qi, Y. (2007). A complex system of small RNAs in the unicellular green alga *Chlamydomonas reinhardtii*. *Genes Dev* 21, 1190-1203.

Zipprich, J.T., Bhattacharyya, S., Mathys, H., and Filipowicz, W. (2009). Importance of the C-terminal domain of the human GW182 protein TNRC6C for translational repression. *RNA*.

Apéndices

La colaboración en otros proyectos de investigación durante el periodo de realización de la tesis doctoral han dado lugar a los siguientes trabajos:

Pernaute B, Canon S, Crespo M, Fernandez-Tresguerres B and Manzanares M. Comparison of extraembryonic expression of *Eomes* and *Cdx2* in pre-gastrulation chick and mouse embryo unveils regulatory changes along evolution. Manuscrito en preparación.

Fernandez-Tresguerres B, Canon S, Pernaute B, Rayon T, Crespo M and Manzanares M. Evolution of the mammalian embryonic pluripotency gene regulatory network. Manuscrito en preparación.

Alonso ME, Pernaute B, Crespo M, Gomez-Skarmeta JL and Manzanares M (2008) Understanding the regulatory genome. Int. J. Dev Biol. *In press*.

Comparison of extraembryonic expression of *Eomes* and *Cdx2* in pre-gastrulation chick and mouse embryo unveils regulatory changes along evolution

Bárbara Pernaute, Susana Cañon, Miguel Crespo¹, Beatriz Fernandez-Tresguerres and Miguel Manzanares*

Department of Cardiovascular Developmental Biology, Centro Nacional de Investigaciones Cardiovasculares (CNIC), Melchor Fernandez Almagro 3, 28029 Madrid, Spain.

¹Present address: Department of Neurology and Neuroscience, Weill Cornell Medical College, New York, NY 10065, USA.

*Corresponding author: Miguel Manzanares (mmanzanares@cnic.es)

Keywords: *Eomes*, *Cdx2*, extraembryonic, area opaca, trophoectoderm, primordial germ cells.

Running title: *Eomes* and *Cdx2* in pre-gastrulation chick embryos

ABSTRACT

In the mouse blastocyst, *Eomes* and *Cdx2* are critical for establishing the trophoectoderm, the precursor of the placenta that is an evolutionary innovation of mammals. To better understand how the trophoectoderm lineage arose during evolution, we examined the expression of their orthologues in the pre-gastrulation chick embryo. While both genes are expressed in extraembryonic tissues of the early chick embryo, their temporal pattern of expression differs from what occurs in mouse, and other genes that show restricted expression in the mouse trophoectoderm are not expressed in extraembryonic regions of the chick. Also unlike the mouse, chick *Eomes* is expressed in primordial germ cells. Finally, despite the apparent similarity of early extraembryonic expression of *Eomes* in chick and mouse, conserved non-coding elements in the *Eomes* genomic region are unable to drive trophoectoderm restricted expression in the mouse blastocyst, but do so in conserved sites of expression such as the forebrain. These results suggest that the appearance of the trophoectoderm in mammals was accompanied by critical changes in the gene regulatory networks controlling extraembryonic development.

INTRODUCTION

The first differentiation event in the mouse embryo is the segregation of embryonic and extraembryonic tissues, which occurs at 3.5 days after fertilization with the differentiation of the extraembryonic trophoectoderm (TE) from the inner cell mass (ICM). The TE is a cell population that gives rise to most of the lineages that form the invasive placenta, and is considered an evolutionary novelty unique to mammals since no homologous structure exists in non-mammalian vertebrates. The segregation of the TE occurs at the blastocyst stage, with the TE forming an outermost layer surrounding the ICM and the blastocyst cavity. The second lineage decision takes place in the late blastocyst, with some ICM cells segregating to form the primitive endoderm (PE), an extraembryonic endoderm cell layer; the rest of the ICM will become the epiblast (Epi), which differentiates into all lineages of the embryo proper (reviewed in (Rossant and Tam, 2009)).

The first events during the development of the early chick embryo differ from the mouse in several respects. The most marked differences are the complete absence of a TE lineage and the positioning of the embryo on top of the nutritive yolk, a structure absent from mammalian embryos, which are nourished via the TE-derived placenta. Nonetheless, very similar to the mouse, one of the earliest morphological differentiation events in the chick embryo is the separation of extraembryonic and embryonic territories (Bellairs and Osmond, 2005). This segregation occurs shortly before the time of laying (EGK-IX-X) (Eyal-Giladi and Kochav, 1976), when two areas can be distinguished in the blastodisc – a thinner central layer, the *area pellucida* (AP), surrounded by a thicker layer, the *area opaca* (AO). The AO gives rise only to extraembryonic structures, while the embryo lineages are specified from the AP. By EGK-X some AP cells delaminate to generate the primary hypoblast beneath the epiblast of the AP, that will in turn originate the extraembryonic endoderm (Eyal-Giladi and Kochav, 1976). The appearance and formation of the different embryonic and extraembryonic populations in the chick embryo is clearly reminiscent of events in the mouse (O'Farrell et al., 2004).

Specification of the TE in the mouse embryo has been studied in detail, and this work has identified some of the main factors in the gene regulatory network that segregates extraembryonic and embryonic lineages at the blastocyst stage. Prior to formation of the blastocyst, the complementary expression patterns and reciprocal repression of the transcription factors *Cdx2* (in the outermost cells) and *Pou5f1* (in the innermost ones) defines the segregation of TE and ICM (Niwa et al., 2005). *Cdx2* then activates *Eomes* expression in the TE, both transcription factors being essential for the specification and maintenance of the TE-derived trophoblast lineage. Conversely, the ICM differentially expresses the pluripotency gene *Nanog* in cells that will form the epiblast, and *Gata4/Gata6* in future PE cells (Rossant and Tam, 2009).

The T-box transcription factor *Eomes* was described as the first pan-mesodermal gene expressed in xenopus (Ryan et al., 1996). Its expression in the primitive streak and role in mesoderm specification during gastrulation are widely conserved with other vertebrates,

including zebrafish, chick and mouse (Bulfone et al., 1999; Russ et al., 2000; Bruce et al., 2003; Arnold et al., 2008). A later expression pattern in the developing brain is also highly conserved in these species. Mouse *Eomes* is expressed specifically in the TE at blastocyst stages and later in the extraembryonic ectoderm, and mutant embryos for this gene die around implantation due to trophoblast defects. This places this transcription factor at a crucial position in the specification and maintenance of the trophoblast lineage, downstream of *Cdx2* (Russ et al., 2000; Strumpf et al., 2005). *Eomes* thus appears to have been recruited during mammalian evolution to the regulatory network that determines the first lineage choice during development and trophoblast fate.

To gain insight into the evolutionary origin of the mammalian trophoblast and of early lineage determination, we analyzed the expression of *Eomes* and *Cdx2* in prestreak and early gastrulating chick embryos. We find that *Eomes* and *Cdx2* are sequentially expressed in extraembryonic regions of the chick embryos. However, these genes are expressed in the reverse order to that seen in the mouse TE, and conserved non-coding elements in the *Eomes* region were unable to drive TE-restricted expression in the mouse blastocyst. Analysis of *Eomes* expression in chick embryos identified a novel and non conserved expression domain in primordial germ cells (PGC). Our results point to fundamental differences in the regulation of extraembryonic fate in chick and mouse, despite morphological and gene expression similarities between them. Future studies will be needed to determine whether *Eomes* has a role in PGC development in chick.

RESULTS AND DISCUSSION

The first lineage decisions in mouse and chick embryos are regulated differently

In the mouse, *Eomes* is maternally expressed and thus can be detected in the oocyte and during the first cleavage stages (McConnell et al., 2005). At the late blastocyst stage *Eomes* expression, under the regulation of *Cdx2* (Strumpf et al., 2005), is restricted to the TE and, after implantation, to its derivative, the extraembryonic ectoderm (Ciruna and Rossant, 1999; Russ et al., 2000). In the gastrulating embryo at E6.5, *Eomes* is also expressed on the posterior side of the embryo – a pattern similar to that seen in xenopus, zebrafish and chick – where it later plays a role in mesoderm specification and epithelium-to-mesenchyme transition (Arnold et al., 2008).

Analysis of the early chick embryo by *in situ* hybridization detected strong expression of *Eomes* at the earliest stage we could examine, corresponding to the freshly laid egg (EGK-X; Fig 1A-C) (Eyal-Giladi and Kochav, 1976). At this stage one can distinguish the extraembryonic AO, and the epiblast and hypoblast of the AP. *Eomes* is expressed in the AO and the hypoblast but is excluded from the epiblast (Fig 1B-C). At Hamburger and Hamilton Stage 1+ (HH1+) (Bellairs and Osmond, 2005), the expression in the AO and hypoblast persist (Fig 1D-F), but *Eomes* is also detected in Koller's sickle and the emerging primitive streak (Fig 1D). Expression in the AO

and hypoblast decays by HH2, at which stage we detected *Eomes* mRNA only in the primitive streak (Fig 1G-I).

Expression of *Eomes* in the chick hypoblast might be related to the role of this gene in endoderm specification in both zebrafish and mouse. However, mouse *Eomes* is not expressed in the PE, the equivalent of the chick hypoblast (O'Farrell et al., 2004), but instead in the anterior primitive streak, from where the definitive endoderm will form (Arnold et al., 2008). In zebrafish it is the maternal *Eomes* product that is localized to the marginal blastomeres, where it induces formation of the endoderm (Bjornson et al., 2005). Maternal expression of *Eomes* has also been reported in mouse (McConnell et al., 2005), although with no spatial restriction. We can not rule out that the high amounts of *Eomes* mRNA detected at EGK-X might represent maternal transcripts, since this stage coincides with the transition from maternal to zygotic gene expression in the chick embryo (Zagris et al., 1998; Elis et al., 2008). Further studies will be required to determine whether a role for *Eomes* in endoderm specification is conserved in the chick embryo and if this is related to its expression in the hypoblast. In this regard, it is interesting that the extraembryonic endoderm contributes to the definitive endoderm in both chick and mouse (Lawson and Schoenwolf, 2003; Kwon et al., 2008).

The early expression of chick *Eomes* in the AO suggests that the similarities between extraembryonic tissues in mouse and chick may extend beyond mere morphological features. We therefore examined the expression of orthologues of other genes that are TE-specific in mouse. *Cdx2* is the first gene known to be specifically expressed in the mouse blastocyst and has been shown to be upstream of *Eomes* during TE specification. *Cdx2*-deficient embryos manifest an earlier TE phenotype than *Eomes*-deficient embryos, and while *Eomes* expression is lost in *Cdx2* knockouts, *Cdx2* expression is intact in *Eomes* knockouts (Strumpf et al., 2005; Ralston and Rossant, 2008). However, in the early chick embryo, and contrary to what we see for *Eomes*, we did not detect *Cdx2* expression at stage EGK-X-XI (Fig 1J), and only by HH1+ we began to detect weak expression in the AO (Fig 1K). By HH2 strong *Cdx2* expression was evident in the AO, at a time when *Eomes* has been downregulated (Fig 1L). At later stages *Cdx2* expression is restricted to the caudal part of the primitive streak (data not shown), in a pattern identical to that previously reported (Marom et al., 1997). According to these results, *Eomes* and *Cdx2* expression domains only overlap during a very restricted period of time in the extraembryonic region of the chick embryo at pre-gastrulation stages (HH1+), when *Cdx2* begins to be expressed and *Eomes* expression is about to decay. This differs markedly from what happens in the mouse blastocyst, and makes it difficult to position *Cdx2* upstream of *Eomes* in the chick embryo. These results thus argue against an overall conservation in amniotes of the gene regulatory networks that control extraembryonic fate.

Given that *Cdx2* is absent from the early extraembryonic lineages of the chick embryo, we examined whether the same was also the case for other mouse TE specific genes. *Bmp4* and *Fgfr2* are expressed in the mouse TE and its derivative the extraembryonic ectoderm (Haffner-Krausz et al., 1999; Lawson et al., 1999; Rossant and Cross, 2001). These factors form part of

the trophoblast-epiblast crosstalk network: *Fgfr2* expressed on trophoblast cells binds *Fgf4* from the epiblast, activating signals that maintain the trophoblast (reviewed in (Rossant and Cross, 2001)), while trophoblast-expressed *Bmp4* maintains the patterning of the epiblast (Fujiwara et al., 2002; Di-Gregorio et al., 2007). Whole mount *in situ* hybridization in early chick embryos revealed that neither *Bmp4* (Fig M-N) nor *Fgfr2* (Fig 1O-P) are expressed until stage HH3, confirming previous reports (Chapman et al., 2002; Lunn et al., 2007): these factors show no expression whatsoever in chick extraembryonic domains.

In both chick and mouse, expression of *Eomes* is initially restricted to extraembryonic lineages, what could indicate a similar regulation of the first lineage decisions in avians and mammals. However, other factors involved in the specification and maintenance of mammalian TE are not expressed in a comparable manner in the early chick extraembryonic domains, and the later onset of *Cdx2* expression in chick appears to indicate that *Eomes* expression in extraembryonic lineages is not regulated in the same manner as in mouse. Moreover, the absence of *Bmp4* and *Fgfr2* suggests that the extraembryonic-epiblast communication network is a novel acquisition in mammals.

***Eomes* is expressed in the primordial germ cells of the chick but not the mouse embryo**

Bulfone et al. (1999) have previously described the expression pattern of chick *Eomes* from stages HH3 to HH28. According to this report, at HH3 *Eomes* is expressed in the anterior-most part of the AP, in the hypoblast and ectoderm anterior and lateral to the PS. This expression decays by HH5 with the regression of the PS, and disappears by HH6. Later in development, at HH25, *Eomes* is expressed in the telencephalic pallium of the developing brain. This pattern, with expression in the gastrulating embryo and later in the central nervous system, is widely conserved in xenopus, zebrafish and mouse (Bulfone et al., 1999; Ciruna and Rossant, 1999; Kimura et al., 1999; Mione et al., 2001; Bachy et al., 2002). In embryos from HH19 to HH24, we detected *Eomes* expression in the previously described territories, such as the telencephalon (Fig 2D-E), but also detected *Eomes* mRNA in a novel and non-conserved domain: the primordial germ cells (PGC).

Eomes expression in PGC becomes apparent by HH4 in the germinal crescent, located at the anterior boundary of the AO and AP (Fig 2A, A'). PGC are large cells that appear dispersed in the germinal crescent and have a very characteristic rounded shape, making them easily distinguishable from other cell populations (Bellairs and Osmond, 2005). At HH5 *Eomes* expression in the primitive streak decreases and becomes restricted to its most anterior portion (Fig 2B), and by HH6 expression is detected anterior to the node along the regressing primitive streak and in the head fold (Fig 2C). At these stages the PGC in the germinal crescent remain positive for *Eomes* (Fig B', C'). Starting at HH10, PGC migrate from the germinal crescent via the vascular system towards the gonads that they reach by HH16 (Bellairs and Osmond, 2005). We detected *Eomes* expression along the genital ridges in wholemount embryos from HH19 to

HH24 (Fig 2D, F; and data not shown), and restriction of this expression to PGC was confirmed in tissue sections of HH22 embryos (Fig 2G).

Expression of *Eomes* in PGC has not been reported in any other vertebrate. We therefore re-examined the expression pattern of *Eomes* in the mouse embryo, focusing on the genital ridges and gonads. PGC are specified in mouse at around E7.0, in the region where the mesoderm is being determined (reviewed in (Matsui and Okamura, 2005)). At this time *Eomes* is highly expressed in mesoderm, making it difficult to ascertain if it is expressed in the PGC. At E10.5 we detected *Eomes* in the previously described domain in the developing brain, but found no expression in the genital ridges (Fig 2H-I). Conversely, at this stage we found clear expression of *Pou5f1*, a well characterized PGC marker (Fig 2K) (Scholer, 1991). Similarly, no expression of *Eomes* was detected in the gonads of E13.5 embryos (Fig 2J), contrasting with the gonadal expression of *Pouf51* at this stage (Fig 2L). The lack of expression of mouse *Eomes* in gonads was confirmed by RT-PCR (data not shown). PGC therefore represent a novel and non-conserved expression domain of *Eomes* in chick.

Several genes are expressed in developing PGC in chick, such as *Vasa*, *Dead end*, *Nanog* and the *Pou5f1*-related gene *PouV* (Tsunekawa et al., 2000; Stebler et al., 2004; Canon et al., 2006; Laval et al., 2007). Of special interest are *Nanog* and *Pou5f1*, because of their pivotal role in specifying the early lineages of the mouse blastocyst and in maintaining embryonic stem cell pluripotency within the epiblast at this stage (Boiani and Scholer, 2005; Niwa, 2007). This places them as repressors of extraembryonic fate in mouse, contrary to *Eomes*, which contributes to TE formation and maintains trophoblast stem cells. Furthermore, mouse Oct4 and Nanog bind to genomic regions in the vicinity of *Eomes*, and are thought to directly repress its transcription (Loh et al., 2006). Accordingly, *Nanog* and *Pou5f1*, but not *Eomes*, are expressed in PGC in mouse (Scholer, 1991; Hatano et al., 2005; Yamaguchi et al., 2005). In contrast, in chick we find *Eomes* expressed together with *Nanog* and *PouV* in this stem cell population. The role of *Eomes* in PGC and its relation to *Nanog* and *PouV* in the chick remain to be elucidated.

Conserved elements in the genomic region of the mouse and chick *Eomes* genes are unable to drive TE-specific reporter expression in blastocysts

Given the common expression of *Eomes* in extraembryonic domains of the mouse blastocyst and the pre-streak chick embryo, we reasoned that these domains might be controlled by conserved regulatory mechanisms. We therefore examined whether non-coding regions surrounding the mouse *Eomes* gene and conserved with chick were able to direct restricted reporter expression in the TE of transgenic mouse blastocysts.

We first compared the *Eomes* genomic regions in human, mouse, opossum and chick, using the gene annotations provided by the Ensembl Genome Browser server (www.ensembl.org) (Fig. 3A). Human *EOMES* is located on chromosome 3, flanked by two intergenic regions of 519 kb (upstream) and 260 kb (downstream). Murine *Eomes* is located on chromosome 9, opossum

Eomes on chromosome 8, and the chick *Eomes* orthologue on chromosome 2. There is conserved synteny among the upstream regions of all species. However, the downstream region found in human, opossum and chick is absent in mouse. This is likely due to an intra-chromosomal rearrangement, resulting in the region downstream of *Eomes* of mouse being syntenic with a region located 10Mb upstream of human *EOMES*. The fact that the downstream region of mouse *Eomes* has suffered a rearrangement, implying a major reduction in the intergenic non-coding region, strongly indicates that most *Eomes* regulatory elements are likely to be located in the upstream region. Recently, a BAC transgenic mouse strain has been described (*Tg(Eomes::GFP)*) that recapitulates *Eomes* expression in the mouse (Kwon and Hadjantonakis, 2007). This BAC construct covers ~186 kb upstream and ~18 kb downstream of *Eomes*, confirming that most if not all regulatory elements are located upstream of *Eomes*, and for that reason we restricted our analysis to that region (Fig 3A, B). Multiple-alignment of sequences from diverse organisms corresponding to the same genomic region is a powerful tool for identifying cis-regulatory elements (Boffelli et al., 2004; de la Calle-Mustienes et al., 2005). This approach is based on the assumption that a high level of conservation in a non-coding region must be due to a functional constraint. We performed multiple-alignment analysis of the non-coding regions flanking *Eomes* in mouse, human, opossum and chick, and selected highly conserved elements (HCE) present in all four species. Five HCE (Eo1 to Eo5) were selected for testing of their enhancer activity in a mouse transient transgenic assay (Fig 3B). None of the HCE tested was able to drive expression of the reporter gene *lacZ* in blastocysts to the level induced by a previously described *Pou5f1* enhancer element, used as positive control (Fig. 3C) (Yeom et al., 1996). Instead, all the HCE drove low-level punctuated expression indistinguishable from a negative control (Fig. 3C; see Experimental Procedures). It is interesting that HCE Eo5 lies within one of the regions detected in a genome-wide chromatin immunoprecipitation assay with p300 designed to identify forebrain enhancers with extremely high accuracy (Visel et al., 2009). When we tested this HCE for its enhancer activity in transgenic embryos at E10.5 we could detect specific expression of the reporter gene *lacZ* in the telencephalon (Fig. 3D). This implies that elements conserved between mouse and chick do indeed control conserved aspects of *Eomes* expression, such as in the forebrain.

The results of our transgenic analysis suggest that the enhancer elements responsible for *Eomes* expression in the TE of the mouse blastocyst are not located in regions highly conserved in chick. This supports the idea that *Eomes* is regulated differently in the extraembryonic tissues of mouse and chick embryos. Nevertheless, we cannot rule out the possibility that conserved binding sites might exist in less conserved regions that fall below the threshold used in our analysis.

Conclusions

By examining the early expression pattern of *Eomes* in early chick embryos we have found surprising similarities as well as differences with its mouse orthologue. Although *Eomes* expression in the chick AO at pre-gastrulation stages seems comparable to that in the TE of the mouse blastocyst, it is not driven by conserved non-coding regions acting as cis-regulatory elements. We cannot exclude the possibility that a common mechanism of *Eomes* expression is mediated by conserved upstream regulatory elements that were missed by our analysis; however, the analysis of *Cdx2* expression in the chick supports the conclusion that the restricted expression of *Eomes* in extraembryonic regions of avians and mammals is established by different mechanisms. In this regards, it is interesting to note that, in the mouse morula, differential cell polarity of outer and inner cells seems to be a critical factor in the establishment of the trophoectoderm and the initial expression of *Cdx2* (Rossant and Tam, 2009). In the case of the pre-gastrulation chick embryo, there is no difference in polarity of cells between area opaca and area pellucida, precluding this being a conserved mechanism for activation of *Cdx2* or *Eomes*.

We have also found that, unexpectedly, chick *Eomes* is expressed in PGC, which does not occur in mouse. Since chick PGC also express *Nanog* and the *Oct4*-related *PouV* genes, and in mouse *Eomes* is repressed by these factors as part of the embryonic pluripotency network, it is tempting to speculate that a change might have occurred in the interrelationship of these genes in the avian and mammalian lineages. This would lend further support to the idea that significant changes in the regulation of *Eomes* expression in the early embryo occurred concomitantly with the appearance of the blastocyst and the early lineage segregation observed in mammals.

EXPERIMENTAL PROCEDURES

Gene Expression Analysis

Whole-mount *in situ* hybridization in chick and mouse embryos was carried out following previously described procedures (Nieto et al., 1996; Ariza-McNaughton and Krumlauf, 2002). The chick *Eomes* probe was generated by amplification of cDNA obtained from whole HH6 embryos with the primers *cEoF* (5'-TGAAGTGAAGATGGCGTTGAG- 3') and *cEoR* (5'-CTTCAGAGAAGCCTGGAGGA-3'). The amplification product was then cloned into pGEM-T Easy Vector for use as the template for reverse transcription. The probes for chick *Cdx2* and *Fgfr2* were obtained from the BBSRC chick EST collection (Boardman et al., 2002). The ESTs selected were ChEST626o23 (*Cdx2*) and ChEST869o7 (*Fgfr2*). The probe for chick *Bmp4* was a gift from Dr. A. Tavares, and mouse *Eomes* and *Pou5f1* probes were gifts from Dr. T. Rodriguez. After whole mount *in situ* hybridization, embryos were cryo-sectioned for microscopy analysis.

Sequence Analysis

Study of the synteny of the *Eomes* genomic context was based on annotations obtained from the Ensembl Genome Browser (www.ensembl.org), from which we also downloaded the genomic sequences used in the comparative analysis. Global multiple alignments were performed using the Vista package (Frazer et al., 2004). Only sequences with at least 70% identity over 100 bp and conserved from human to chicken were categorized as highly conserved elements (HCE).

Transgenic Analysis

HCE selected according to the global multiple alignment analysis were amplified using four selected BACs as templates: RP23-478D17, RP24-77J5, RP23-272A7, RP23-235G22. These BACs cover the whole intergenic region flanking mouse *Eomes* and were obtained from the BACPAC Resources Center (<http://bacpac.chori.org/>). Primers used for PCR and the lengths of corresponding amplified fragments were as follows. *Eo1F*: 5'-GCAGGGAGTGCTTCTGTTTATT-3', *Eo1R*: 5'-ACCCCTTGAGTTGAGTCATAA-3' (2072 bp); *Eo2F*: 5'-TGGGGTGTATATAAGCAGTCCTC-3', *Eo2R*: 5'-TCTGTCTGTCTGTCTATCATCAACC-3' (1402 bp); *Eo3F*: 5'-AGAGGACTTGGGTTTAGTTTCCA-3', *Eo3R*: 5'-GACCCAGGGAACCTTTCTTA-3' (6653 bp); *Eo4F*: 5'-GCAGTACACACACGCATGAA-3', *Eo4R*: 5'-TCATTTCCCAGGATTCTTCG-3' (1402 bp); *Eo5F*: 5'-ATGCCCACAACACACACAGA-3', *Eo5R*: 5'-TGTTACTGGGGCTTGGATGT-3' (3658 bp). As a positive control we amplified the previously described *Pou5f1* distal enhancer element (Yeom et al., 1996), using BAC RP23-152G18 as template (<http://bacpac.chori.org/>); the primers for PCR were *Pou5f1DE-F*: 5'-AGCGGCCGCCTCTGCTACATGTAAATTTGTCT-3' and *Pou5f1DE-R*: 5'-AGCGGCCGCCTAAACAAGTACTCAACCCTTGAA-3' (3368 bp). Each fragment was subcloned in pGEM-T Easy Vector and then excised and cloned into a modified pBluescript vector (Yee and Rigby, 1993) containing the *lacZ* reporter gene under the control of the human beta-globin minimal promoter and with an SV40 polyadenylation signal. Constructs were linearized and plasmid sequences removed prior to micro-injection.

For generation of transient transgenics, females from a F1 C57/CBA background were super-ovulated to obtain fertilized oocytes (Nagy et al., 2003). Each construct was micro-injected into the pronucleus of fertilized oocytes at 0.5 dpc and a concentration of 4ng/ul. Micro-injected oocytes were cultured in microdrops of M16 medium (SIGMA) covered with mineral oil (SIGMA) at 37°C, 5% CO₂ until blastocyst stage. For *lacZ* staining, blastocysts were fixed in buffer containing 1% formaldehyde, 0.2% glutaraldehyde, 2mM MgCl₂, 5mM EGTA, and 0.02% Igepal for 5 min at room temperature. After washes in PBS, blastocysts were transferred to X-Gal staining solution for 24 hours at room temperature in the dark. A minimum of 50 blastocysts were used per construct to calculate the percentage of *lacZ* positive embryos. When using the empty vector containing only the minimal promoter and the *lacZ* reporter as a negative control, we routinely obtain low-level punctuated *lacZ* expression in approximately 4% of blastocysts.

The embryos at E10.5 were fixed for 30 minutes at 4°C and then washed several times in PBS-0.02% Igepal at room temperature prior to staining.

ACKNOWLEDGEMENTS

We wish to thank Ana Tavares and Tristan Rodriguez for in situ probes, Robb Krumlauf for reporter constructs, the Transgenic Unit at CNIC for generating embryos, Eva Alonso and Cristina Arias for fruitful discussions, and Simon Bartlett for critical reading and editing of the manuscript. This work was supported by grants from the Spanish Government (BFU2008-00838 and CONSOLIDER-25120), the Regional Government of Madrid (CAM S-SAL-0190-2006), and the ProCNIC Foundation.

REFERENCES

- Ariza-McNaughton L, Krumlauf R. 2002. Non-radioactive in situ hybridization: simplified procedures for use in whole-mounts of mouse and chick embryos. *Int Rev Neurobiol* 47:239-250.
- Arnold SJ, Hofmann UK, Bikoff EK, Robertson EJ. 2008. Pivotal roles for eomesodermin during axis formation, epithelium-to-mesenchyme transition and endoderm specification in the mouse. *Development* 135:501-511.
- Bachy I, Berthon J, Retaux S. 2002. Defining pallial and subpallial divisions in the developing *Xenopus* forebrain. *Mech Dev* 117:163-172.
- Bellairs R, Osmond M. 2005. *The Atlas of Chick Development*. London, UK: Elsevier Academic Press.
- Bjornson CR, Griffin KJ, Farr GH, 3rd, Terashima A, Himeda C, Kikuchi Y, Kimelman D. 2005. Eomesodermin is a localized maternal determinant required for endoderm induction in zebrafish. *Dev Cell* 9:523-533.
- Boardman PE, Sanz-Ezquerro J, Overton IM, Burt DW, Bosch E, Fong WT, Tickle C, Brown WR, Wilson SA, Hubbard SJ. 2002. A comprehensive collection of chicken cDNAs. *Curr Biol* 12:1965-1969.
- Boffelli D, Nobrega MA, Rubin EM. 2004. Comparative genomics at the vertebrate extremes. *Nat Rev Genet* 5:456-465.
- Boiani M, Scholer HR. 2005. Regulatory networks in embryo-derived pluripotent stem cells. *Nat Rev Mol Cell Biol* 6:872-884.
- Bruce AE, Howley C, Zhou Y, Vickers SL, Silver LM, King ML, Ho RK. 2003. The maternally expressed zebrafish T-box gene eomesodermin regulates organizer formation. *Development* 130:5503-5517.
- Bulfone A, Martinez S, Marigo V, Campanella M, Basile A, Quaderi N, Gattuso C, Rubenstein JL, Ballabio A. 1999. Expression pattern of the Tbr2 (Eomesodermin) gene during mouse and chick brain development. *Mech Dev* 84:133-138.
- Canon S, Herranz C, Manzanares M. 2006. Germ cell restricted expression of chick Nanog. *Dev Dyn* 235:2889-2894.
- Ciruna BG, Rossant J. 1999. Expression of the T-box gene Eomesodermin during early mouse development. *Mech Dev* 81:199-203.
- Chapman SC, Schubert FR, Schoenwolf GC, Lumsden A. 2002. Analysis of spatial and temporal gene expression patterns in blastula and gastrula stage chick embryos. *Dev Biol* 245:187-199.
- de la Calle-Mustienes E, Feijoo CG, Manzanares M, Tena JJ, Rodriguez-Seguel E, Letizia A, Allende ML, Gomez-Skarmeta JL. 2005. A functional survey of the enhancer activity of conserved non-coding sequences from vertebrate Iroquois cluster gene deserts. *Genome Res* 15:1061-1072.
- Di-Gregorio A, Sancho M, Stuckey DW, Crompton LA, Godwin J, Mishina Y, Rodriguez TA. 2007. BMP signalling inhibits premature neural differentiation in the mouse embryo. *Development* 134:3359-3369.
- Elis S, Batellier F, Couty I, Balzergue S, Martin-Magniette ML, Monget P, Blesbois E, Govoroun MS. 2008. Search for the genes involved in oocyte maturation and early embryo development in the hen. *BMC Genomics* 9:110.
- Eyal-Giladi H, Kochav S. 1976. From cleavage to primitive streak formation: a complementary normal table and a new look at the first stages of the development of the chick. I. General morphology. *Dev Biol* 49:321-337.
- Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. 2004. VISTA: computational tools for comparative genomics. *Nucleic Acids Res* 32:W273-279.
- Fujiwara T, Dehart DB, Sulik KK, Hogan BL. 2002. Distinct requirements for extra-embryonic and embryonic bone morphogenetic protein 4 in the formation of the node and primitive streak and coordination of left-right asymmetry in the mouse. *Development* 129:4685-4696.
- Haffner-Krausz R, Gorivodsky M, Chen Y, Lonai P. 1999. Expression of Fgfr2 in the early mouse embryo indicates its involvement in preimplantation development. *Mech Dev* 85:167-172.
- Hatano SY, Tada M, Kimura H, Yamaguchi S, Kono T, Nakano T, Suemori H, Nakatsuji N, Tada T. 2005. Pluripotential competence of cells associated with Nanog activity. *Mech Dev* 122:67-79.
- Kimura N, Nakashima K, Ueno M, Kiyama H, Taga T. 1999. A novel mammalian T-box-containing gene, Tbr2, expressed in mouse developing brain. *Brain Res Dev Brain Res* 115:183-193.

- Kwon GS, Hadjantonakis AK. 2007. Eomes::GFP-a tool for live imaging cells of the trophoblast, primitive streak, and telencephalon in the mouse embryo. *Genesis* 45:208-217.
- Kwon GS, Viotti M, Hadjantonakis AK. 2008. The endoderm of the mouse embryo arises by dynamic widespread intercalation of embryonic and extraembryonic lineages. *Dev Cell* 15:509-520.
- Lavial F, Acloque H, Bertocchini F, Macleod DJ, Boast S, Bachelard E, Montillet G, Thenot S, Sang HM, Stern CD, Samarut J, Pain B. 2007. The Oct4 homologue PouV and Nanog regulate pluripotency in chicken embryonic stem cells. *Development* 134:3549-3563.
- Lawson A, Schoenwolf GC. 2003. Epiblast and primitive-streak origins of the endoderm in the gastrulating chick embryo. *Development* 130:3491-3501.
- Lawson KA, Dunn NR, Roelen BA, Zeinstra LM, Davis AM, Wright CV, Korving JP, Hogan BL. 1999. Bmp4 is required for the generation of primordial germ cells in the mouse embryo. *Genes Dev* 13:424-436.
- Loh YH, Wu Q, Chew JL, Vega VB, Zhang W, Chen X, Bourque G, George J, Leong B, Liu J, Wong KY, Sung KW, Lee CW, Zhao XD, Chiu KP, Lipovich L, Kuznetsov VA, Robson P, Stanton LW, Wei CL, Ruan Y, Lim B, Ng HH. 2006. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet* 38:431-440.
- Lunn JS, Fishwick KJ, Halley PA, Storey KG. 2007. A spatial and temporal map of FGF/Erk1/2 activity and response repertoires in the early chick embryo. *Dev Biol* 302:536-552.
- Marom K, Shapira E, Fainsod A. 1997. The chicken caudal genes establish an anterior-posterior gradient by partially overlapping temporal and spatial patterns of expression. *Mech Dev* 64:41-52.
- Matsui Y, Okamura D. 2005. Mechanisms of germ-cell specification in mouse embryos. *Bioessays* 27:136-143.
- McConnell J, Petrie L, Stennard F, Ryan K, Nichols J. 2005. Eomesodermin is expressed in mouse oocytes and pre-implantation embryos. *Mol Reprod Dev* 71:399-404.
- Mione M, Shanmugalingam S, Kimelman D, Griffin K. 2001. Overlapping expression of zebrafish T-brain-1 and eomesodermin during forebrain development. *Mech Dev* 100:93-97.
- Nagy A, Gertszensten M, Vintersten K, Behringer R. 2003. Manipulating the mouse embryo: a laboratory manual. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
- Nieto MA, Patel K, Wilkinson DG. 1996. In situ hybridization analysis of chick embryos in whole mount and tissue sections. *Methods Cell Biol* 51:219-235.
- Niwa H. 2007. How is pluripotency determined and maintained? *Development* 134:635-646.
- Niwa H, Toyooka Y, Shimosato D, Strumpf D, Takahashi K, Yagi R, Rossant J. 2005. Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* 123:917-929.
- O'Farrell PH, Stumpff J, Su TT. 2004. Embryonic cleavage cycles: how is a mouse like a fly? *Curr Biol* 14:R35-45.
- Ralston A, Rossant J. 2008. Cdx2 acts downstream of cell polarization to cell-autonomously promote trophectoderm fate in the early mouse embryo. *Dev Biol* 313:614-629.
- Rossant J, Cross JC. 2001. Placental development: lessons from mouse mutants. *Nat Rev Genet* 2:538-548.
- Rossant J, Tam PP. 2009. Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. *Development* 136:701-713.
- Russ AP, Wattler S, Colledge WH, Aparicio SA, Carlton MB, Pearce JJ, Barton SC, Surani MA, Ryan K, Nehls MC, Wilson V, Evans MJ. 2000. Eomesodermin is required for mouse trophoblast development and mesoderm formation. *Nature* 404:95-99.
- Ryan K, Garrett N, Mitchell A, Gurdon JB. 1996. Eomesodermin, a key early gene in *Xenopus* mesoderm differentiation. *Cell* 87:989-1000.
- Scholer HR. 1991. Octamania: the POU factors in murine development. *Trends Genet* 7:323-329.
- Stebler J, Spieler D, Slanchev K, Molyneaux KA, Richter U, Cojocaru V, Tarabykin V, Wylie C, Kessel M, Raz E. 2004. Primordial germ cell migration in the chick and mouse embryo: the role of the chemokine SDF-1/CXCL12. *Dev Biol* 272:351-361.
- Strumpf D, Mao CA, Yamanaka Y, Ralston A, Chawengsaksophak K, Beck F, Rossant J. 2005. Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. *Development* 132:2093-2102.

- Tsunekawa N, Naito M, Sakai Y, Nishida T, Noce T. 2000. Isolation of chicken vasa homolog gene and tracing the origin of primordial germ cells. *Development* 127:2741-2750.
- Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, Afzal V, Ren B, Rubin EM, Pennacchio LA. 2009. ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* 457:854-858.
- Yamaguchi S, Kimura H, Tada M, Nakatsuji N, Tada T. 2005. Nanog expression in mouse germ cell development. *Gene Expr Patterns* 5:639-646.
- Yee SP, Rigby PW. 1993. The regulation of myogenin gene expression during the embryonic development of the mouse. *Genes Dev* 7:1277-1289.
- Yeom YI, Fuhrmann G, Ovitt CE, Brehm A, Ohbo K, Gross M, Hubner K, Scholer HR. 1996. Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. *Development* 122:881-894.
- Zagris N, Kalantzis K, Guialis A. 1998. Activation of embryonic genome in chick. *Zygote* 6:227-231.

FIGURE LEGENDS

Fig. 1: Expression of trophoblast markers in prestreak chick embryos. (A-I) *Eomes* expression is detected from stage EGK-X to the beginning of primitive streak formation. At EGK-X (A-C) *Eomes* is expressed in the *area opaca* (B) and the hypoblast of the *area pellucida*, whereas it is excluded from the epiblast. By HH1+ (D-F) expression is maintained in the *area opaca* (E) and hypoblast of the *area pellucida*, and also begins in the Köller's sickle (D). With the beginning of primitive streak formation (HH2, G-I) *Eomes* expression disappears from the *area opaca* (G-H) and becomes restricted to the primitive streak (I). *Cdx2* expression (**J-L**) is not detected until HH1+, when it begins to be weakly expressed in the *area opaca* (K); its expression only becomes strong in the *area opaca* by HH2 (L). *Bmp4* (**M-N**) and *Fgfr2* (**O-P**) are not expressed in extraembryonic structures at prestreak stages (M, O) and their first appearance is coincident with the emergence the primitive streak at HH3+ (N, P). AO: area opaca; AP: area pellucida; Epi: epiblast; Hypo: hypoblast; PS: primitive streak. Horizontal lines in A, D and G indicate the plane of the sections shown in adjacent panels.

Fig. 2: *Eomes* is expressed in primordial germ cells in the chick. (A-C) At HH4, *Eomes* expression is detected in the anterior half of the embryo (A) and in scattered cells (PGC) in the boundary between the anterior area opaca and area pellucida (A': magnification of the square in A). By HH5 (B) the expression in the epiblast is restricted to the PGC and the anterior-most portion of the primitive streak (B': magnification of the square in B; arrowhead indicates PGC). With the regression of the primitive streak (HH6, C) epiblast expression is seen only around the node and in the head fold, and the signal in the PGC along the germinal crescent is very apparent (C': magnification of the square in C). (**D-G**) At later stages *Eomes* is expressed in the forebrain of the chick (white arrow in D; E) and along the genital ridges (black arrow in D; F: magnification of the chick in D). The PGC expressing *Eomes* along the genital ridge are shown in the section in G. (**H-L**) In mouse, *Eomes* is detected in a conserved pattern in the forebrain of the E10.5 embryo (black arrow in H) but was not detected in the genital ridges (red arrow in H; I: magnification of the embryo in H) or in the gonads at E13.5 (J). The PGC marker *Pou5f1* is clearly detected in these territories (K, expression in PGC along the genital ridge of the E10.5 embryo; L, expression in E13.5 gonads). Ovaries are shown on the left and testis on the right (J, L).

Fig. 3: Analysis of *Eomes* regulation in the mouse blastocyst. (A) The genomic region containing *Eomes* is conserved between mammals and chick with the exception of the region downstream of mouse *Eomes*, which has undergone rearrangement and lost synteny with respect to other vertebrates. The region covered by the BAC used to create the mouse transgenic line *Tg(Eomes::GFP)*, which recapitulates *Eomes* expression (Kwon and Hadjantonakis, 2007), is indicated. (**B**) This region was used in a global multiple alignment

analysis, with the mouse as base organism, to identify elements conserved in all species examined (highly conserved elements: HCE). The five HCE identified (Eo1 to Eo5) are highlighted blue. **(C)** Eo1 to Eo5 were tested for their enhancer activity in blastocysts, and none was able to drive expression of *LacZ* to a level comparable to that of the previously described *Pou5f1* enhancer (Pou5f1DE). Scattered dots were detected inside cells of HCE transgenic and control blastocysts, in a pattern like that shown for Eo1. The table shows the percentage of *lacZ* positive blastocysts obtained for each construct. **(D)** Conversely we could detect enhancer activity for Eo 5, that drives expression of the transgene to the forebrain (arrow head), an evolutionarily conserved domain of *Eomes* expression among vertebrates.

Figure 1

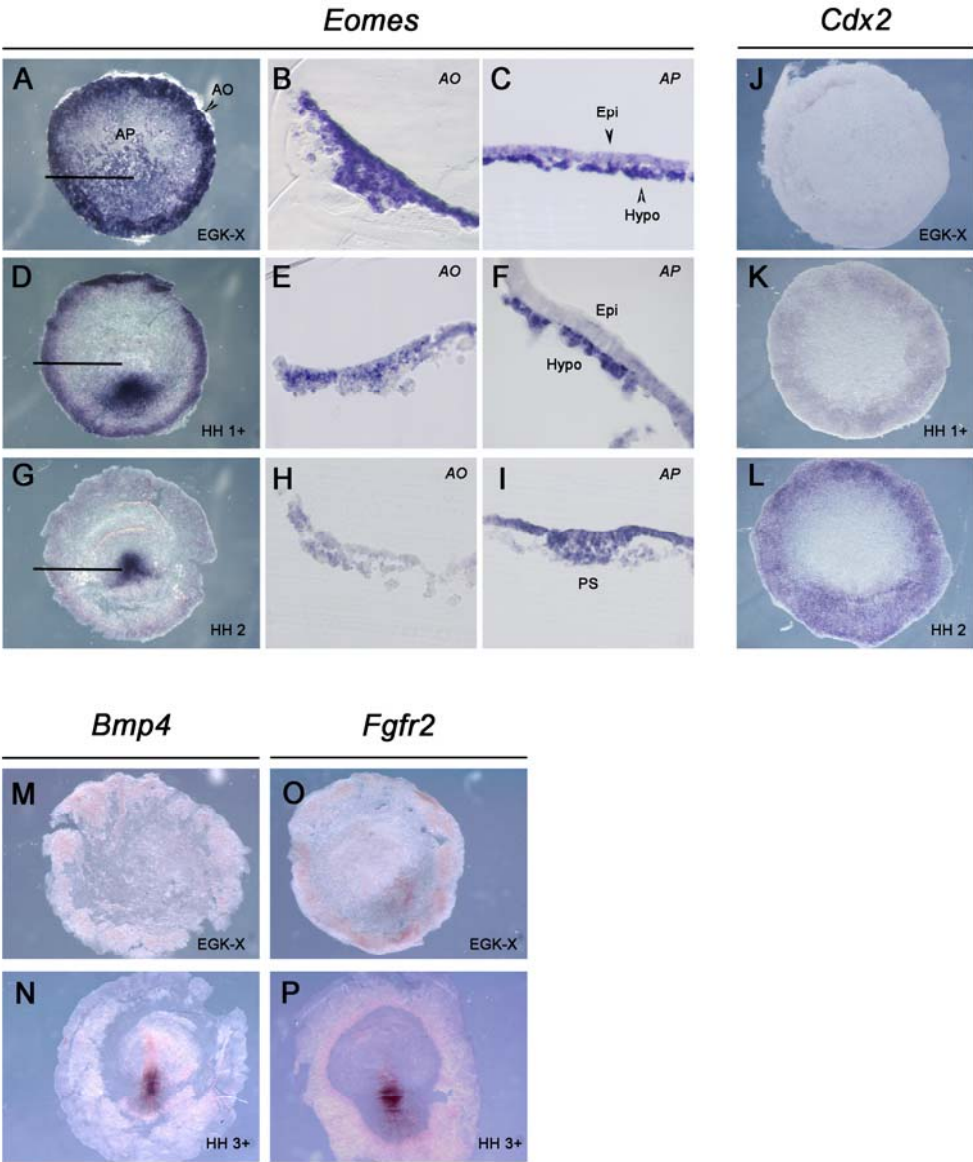


Figure 2

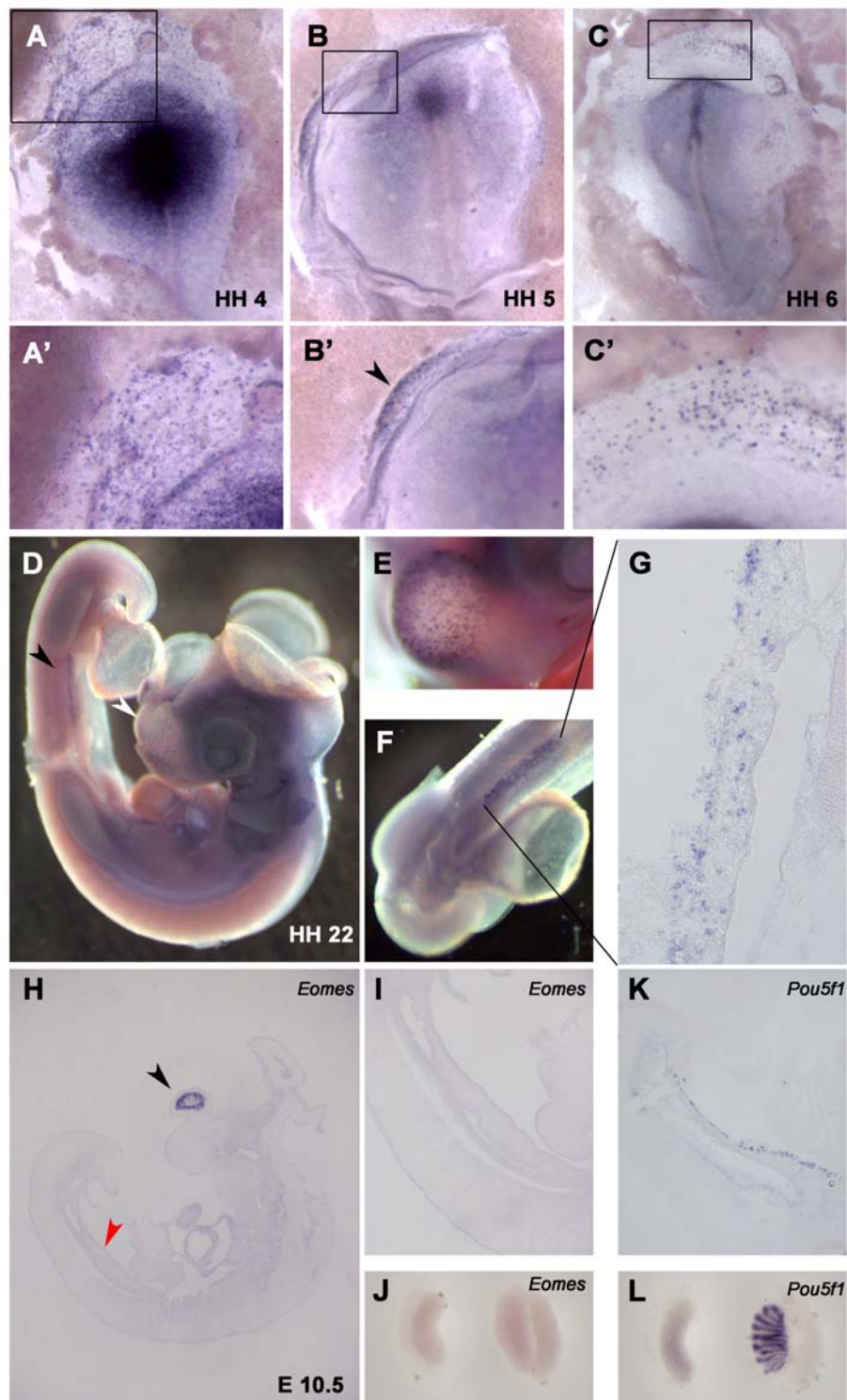
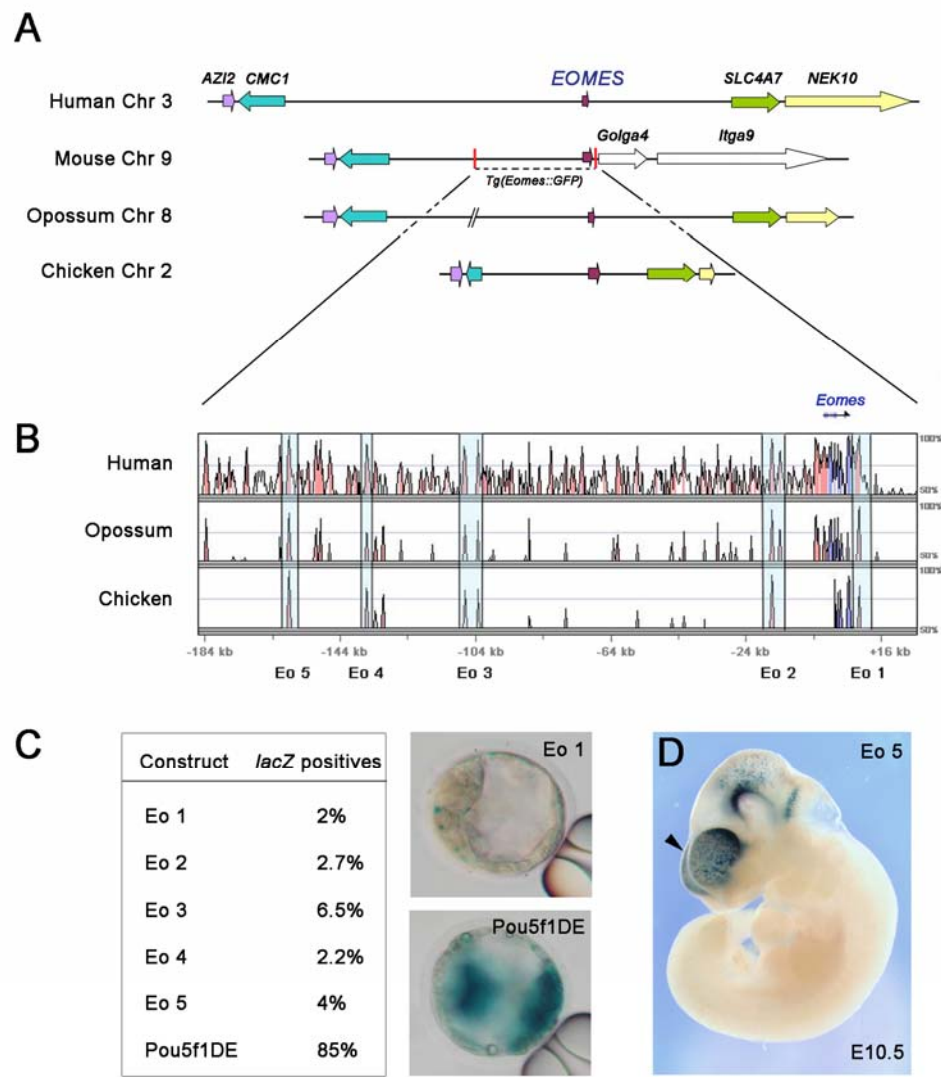


Figure 3



Evolution of the mammalian embryonic pluripotency gene regulatory network

Beatriz Fernandez-Tresguerres¹, Susana Cañon¹, Barbara Pernaute, Teresa Rayon, Miguel Crespo² and Miguel Manzanares³

Department of Cardiovascular Developmental Biology, Centro Nacional de Investigaciones Cardiovasculares (CNIC), Melchor Fernandez Almagro 3, 28029 Madrid, Spain.

¹These authors contributed equally to this work.

²Present address: Department of Neurology and Neuroscience, Weill Cornell Medical College, New York, NY 10065, USA.

³To whom correspondence should be addressed (mmanzanares@cnic.es)

Running title: Evolution of embryonic pluripotency

ABSTRACT

Embryonic pluripotency in the mouse is established and maintained by a gene regulatory network under the control of a core set of transcription factors that include Oct4, Sox2 and Nanog. While this network is largely conserved in eutherian mammals, very little information is available regarding its evolutionary conservation in other vertebrates.

We have performed a detailed comparative analysis of the embryonic pluripotency network between mouse and chick by means of expression analysis in the pre-gastrulation chick embryo, genomic comparisons, and functional assays in ES cells and blastocysts of pluripotency-related regulatory elements. We find that multiple components of the network are either novel to mammals or have acquired novel expression domains in early developmental stages of the mouse. We also find that the downstream action of the mouse core pluripotency factors is largely mediated by genomic sequence elements that are not conserved with chick. In the case of *Sox2* and *Fgf4*, we find that elements driving expression in embryonic pluripotent cells have evolved by a small number of nucleotide changes that create novel binding sites for core factors.

In conclusion, our results show that the network in charge of embryonic pluripotency is an evolutionary novelty of mammals that can be related to an extended period in mammalian development compared to other vertebrates where embryonic cells need to be maintained in an undetermined state prior to engaging in early differentiation events.

INTRODUCTION

Embryonic pluripotency is an essential property of a small group of cells of the mammalian blastocyst, by which they are transiently kept in an indeterminate, uncommitted state. This condition is a consequence of the earliest differentiation events that take place in the pre-implantation embryo. The first lineage decision produces the inner cell mass (ICM) and the trophoectoderm (TE). The TE produces most of the extra-embryonic structures, such as the placenta. In the second lineage choice, the ICM gives rise to the epiblast (EPI), which will generate the embryo proper, and the primitive endoderm (PE), another extra-embryonic tissue (Yamanaka et al. 2006).

Cells in the ICM and the EPI retain the full potential to develop into all embryonic tissues and germ layers, and are the source of embryonic stem (ES) cells, which have the capacity of indefinite self-renewal and maintenance of pluripotency in tissue culture. Pluripotency results from the expression of a small network of transcription factors that actively maintains the undetermined state and at the same time represses the differentiation programme (Boiani and Scholer 2005; Niwa 2007). The core members of this network are the products of the *Oct4* (official symbol *Pou5f1*), *Nanog* and *Sox2* genes (Silva and Smith 2008). These factors act together through auto- and cross-regulatory interactions and also through direct and overlapping binding to multiple locations across the genome, where they regulate downstream target genes (Boyer et al. 2005; Loh et al. 2006; Marson et al. 2008). These transcription factors lie at the core of an extended gene regulatory network (GRN) which includes other transcription factors as well as epigenetic modifiers and signalling molecules; the coordinated activities of this GRN is ultimately responsible for the pluripotent state of ES cells (Chen et al. 2008; Kim et al. 2008).

Certain features of early lineage determination in the mammalian blastocyst can be considered evolutionary novelties, in that they do not appear in other vertebrates. For example, non-mammalian vertebrates have no homologue of the TE, and the placenta, the TE's main derivative, is a clear mammalian innovation. This raises the question of the extent to which genetic control of these early phases of mammalian development, in particular the establishment and maintenance of embryonic pluripotency, is conserved in other vertebrates.

So far little is known about the evolutionary conservation of the embryonic pluripotency GRN (EP-GRN), being limited to the description of non-mammalian vertebrate homologues of the core mammalian EP transcription factors *Oct4* and *Nanog* (Burgess et al. 2002; Bachvarova et al. 2004; Canon et al. 2006; Morrison and Brickman 2006; Laval et al. 2007). While it appears that the biochemical properties of these factors are conserved to some extent between mouse and other vertebrates such as *Xenopus* (Morrison and Brickman 2006) or chick (Laval et al. 2007), other aspects, such as the territories of early expression, are not (Canon et al. 2006).

In order to extend this comparison, we have examined the pre-gastrulation chick embryo for the expression of homologues of important mouse EP-GRN genes. We report that the expression pattern of these homologues is incompatible with a role in establishing embryonic pluripotency

in the early chick embryo, suggesting that these genes must have been newly recruited to this role in mammals. Genomic analysis shows that while some EP-GRN genes are specific to mammals, core pluripotency factors, including *Oct4* which has a complex duplicative history in vertebrates (Niwa et al. 2008), are present throughout amniotes. Finally, we show that the genomic regions that show DNA binding of the core EP transcription factors are poorly conserved between mammals and chick, and that key EP regulatory elements have appeared de novo in mammals. Thus, our results indicate that the EP-GRN is a mammalian novelty that has arisen from co-option, duplication and the appearance of novel genes and regulatory interactions.

RESULTS

Orthologues of mammalian EP-GRN genes are not expressed in the pre-gastrulation chick embryo

We examined the expression of the chick orthologues of several genes part of the EP-GRN at pre-gastrulation stages. These genes include *Sox2*, one of the core factors of the EP-GRN; *FoxD3*, which is required downstream of *Oct4/Sox2* and *Nanog* to maintain stem cells in the embryo (Hanna et al. 2002); and *Fgf4* and *Nodal*, which are expressed in the ICM of the blastocyst (Niswander and Martin 1992; Takaoka et al. 2006) and encode signalling molecules required for proper specification and differentiation of the blastocyst lineages (Boiani and Scholer 2005). We also analyzed the polycomb group gene *Phc1*, because it is a target of the core EP factors (Boyer et al. 2005; Loh et al. 2006), is expressed in ES cells, and maps close to the *Nanog-Gdf3-Dppa3* pluripotency gene cluster (Clark et al. 2004; Canon et al. 2006).

None of these genes are expressed in the pre-gastrulation chick embryo (Stage EGK-X; Fig. 1), contrary to *Pou2*-related (Niwa et al. 2008) that we find expressed throughout the embryo at EGK-X and later strongly upregulated in the area pelucida, in a similar pattern to that recently described (Laval et al. 2007). We detect onset of expression for this set of genes only at later stages in the patterns and domains that in some cases have been previously described (Fig. 1): *Sox2* in the neural plate at HH4 (Rex et al. 1997); *FoxD3* in the node at HH3, earlier than reported (Kos et al. 2001); *Phc1* in the neural tube at HH9, a pattern similar to its late expression in the mouse embryo (Tomotsune et al. 2000); *Fgf4* in the primitive streak at HH4 (Shamim and Mason 1999); and *Nodal* in the node and primitive streak at HH3 (Chapman et al. 2002). Together with our earlier description of *Nanog* expression in chick (Canon et al. 2006), this result shows that orthologues of many genes involved in establishing and maintaining embryonic pluripotency and early lineages in the mouse are not expressed in equivalent territories of the early chick embryo, such as the epiblast.

Several EP-GRN genes are specific to mammals

Our search for chick orthologues of mouse EP-GRN genes showed that in many cases they are not present in the available genomic and transcriptional databases. This could be due simply to lack of coverage of genomic and EST sequences, and we therefore examined these cases in detail to determine if these genes were unique to the mammalian lineage, as is the case for the pluripotency-associated genes *Gdf3* and *Dppa3*, which map adjacent to *Nanog* (Canon et al. 2006).

The chromosomal regions surrounding the mouse genes *Utf1* (Okuda et al. 1998), *Tex19.1*, *Dppa2*, *Dppa4* and *Dppa6* conserves synteny with the chick genome, but none of these genes is present in chick, although their immediate neighbours are (Fig. 2). Extensive searches found no evidence for orthologues of these genes in other genomic positions or in other databases. *Tex19.1* and its tandem duplicate *Tex19.2* have been previously described as mammal-specific (Kuntz et al. 2008), although the comparison was with teleost fishes and invertebrates and did not include chicken. It is noteworthy that these genes are adjacent to another pair of tandem duplicates present in mouse but not in chick, *Sectm1a* and *Sectm1b* (Fig. 2B). These genes belong to the Ig superfamily and are closely related to the neighbouring *Cd7* gene, which is conserved in chick. This indicates that this region has undergone extensive duplication and gene gain during mammalian evolution. *Dppa5* is flanked by *Omt2b* (*Oocyte maturation beta*) and *Ooep* (*Oocyte expressed protein homolog*), two other mammal-specific genes (Fig. 2C). These genes are expressed in a very similar pattern to *Dppa5* in the embryo and ES cells (Popova and Morris 2004; Tanaka et al. 2006), and furthermore *Ooep* and *Dppa5* belong to the same family (Pierre et al. 2007). *Dppa2* and *Dppa4* are tandem duplicated genes expressed in the ICM and later in primordial germ cells (Maldonado-Saldivia et al. 2007), and interestingly forced expression of *Dppa4* in ES cells drives them to a primitive ectoderm lineage (Masaki et al. 2007). The *Dppa2/4* duplication must have occurred late in mammalian evolution, since the opossum contains only one *Dppa2/4* gene (not shown). The neighbouring *Morc1* gene (*microorchidia*) is involved in spermatogenesis (Inoue et al. 1999), and also has no orthologue in chick (Fig. 2D). Unigene-based EST expression profiling (Sayers et al. 2008) shows that this gene is also expressed in the pre-implantation mouse embryo.

Other EP-GRN genes not found in the chicken genome are not located in syntenic regions conserved between mouse and chick. Therefore the only support for their being specific to mammals is that no orthologues have been found in sequence databases of chicken or other non-mammalian species. Such is the case of *Dppa1* (Bortvin et al. 2003) and the *miR-290* cluster (Houbaviy et al. 2003; Houbaviy et al. 2005), which are nonetheless considered as specific to eutherians. For other genes, additional evidence is consistent with their being specific to mammals. For example, *Rex1* (*Zfp42*), which is a retrotransposed copy of *YY1* and appears to be specific to placental mammals (Kim et al. 2007), lies in an extremely gene-poor region of mouse chromosome 8 that is enriched in mammal-specific genes such as *Triml1*, *Adam26a*, *Adam26b* and *Adam34*. The databases contain no chick orthologue of *Nac1* (Kim et al. 2008) but orthologues are found in fish and *Xenopus*. *Nac1* is located in a region of mouse chromosome 8 (40 Mb from *Rex1*) where all surrounding genes show the same pattern of

conservation. This could indicate that the specific genomic region containing *Nac1* has been deleted in avian evolution, in line with recent evidence showing a high level of gene loss in avians compared with other vertebrates (Hughes and Friedman 2008).

This analysis thus reveals that many genes with pivotal roles in embryonic pluripotency in the mouse are not present in the chick, suggesting that they may have arisen specifically in the mammalian lineage in connection with the establishment of the EP-GRN.

The duplication of *Oct4* predates the emergence of amniotes

The situation described above resembles recent results concerning the evolution of the *Oct4* gene (Niwa et al. 2008). A number of labs had previously identified *Oct4*-related genes from various non-mammalian vertebrates (Burgess et al. 2002; Bachvarova et al. 2004; Morrison and Brickman 2006; Laval et al. 2007), claiming that such genes were the true homologues of mammalian *Oct4* based on sequence similarity, early embryonic expression, conservation of synteny (Burgess et al. 2002; Morrison and Brickman 2006), or functional conservation in tissue culture assays (Morrison and Brickman 2006; Laval et al. 2007). However, recent work has claimed that the *Pou5* class was represented by a single gene in the early vertebrate lineage, which was duplicated before the mammalian radiation giving rise to a second *Pou5* gene as is the situation nowadays in non-eutherian mammals. Subsequently, one of the copies was lost in eutherian mammals resulting in the single *Oct4* gene present in mouse or humans (Niwa et al. 2008). Therefore the pluripotency gene *Oct4* is not the orthologue of chicken *Pou2*-related and other vertebrate *Pou5* genes, but a paralogue. Previous claims of conserved synteny between zebrafish and *Xenopus* *Pou5* genes and mammalian *Oct4* (Burgess et al. 2002; Morrison and Brickman 2006) were surely confused by the fact that there is linkage of both *Pou2*-related and *Oct4* to members of the chloride intercellular channel family (*Clic3* and *Clic1*, respectively; Fig. S1B).

We could not find orthologues of genes surrounding mouse *Oct4* in the chicken genome, trace archives or EST databases. Because such genes are present in other vertebrates, most probably this region has been specifically lost in the avian lineage (Fig. S1A). We found the same situation when we examined the recently assembled genome of another bird, the zebrafinch (*Taeniopygia guttata*; not shown). This view is supported by the fact that mammalian *Oct4* is located in the extended MHC Class III region, which is conserved with *Xenopus* but has been lost in chicken and other avians (Deakin et al. 2006). Such finding means that the exact time of duplication of the ancestral vertebrate *Pou5* gene cannot be assigned unambiguously to the base of the mammalian lineage, as we can not rule out that it occurred before divergence of the avian lineage and that it was subsequently lost as part of the whole region.

Such hypothesis was confirmed when we analyzed the genome of the green anole lizard (*Anolis carolinensis*). We found a single predicted transcript more closely related to *Oct4* than to other *Pou2-r* genes (not shown) that mapped to a region that is in conserved synteny with mammalian

Oct4 (Fig. S1B). We failed to find a second *Pou5* like gene, but we believe this is due to the lack of genome coverage in the region. We were able to find linkage of lizard *Abca2* to *Clic3* genes and on the other hand the orthologue of *Npdc1*, but on genomic scaffolds that did not extend further into the region that supposedly would contain *Pou2-r* (Fig. S1B). Nevertheless, the existence of an unambiguous *Oct4* orthologue in the lizard genome suggest that the duplication of the ancestral *Pou5* gene occurred before the divergence of amniotes, and therefore the *Oct4* orthologue was lost in the avian lineage. These results discard any causal relationship between the duplication of *Pou5* genes and the appearance of novel regulatory interaction in the early mammalian embryo, as has been suggested (Niwa et al. 2008).

Core EP-GRN binding cassettes have been acquired in the mammalian lineage

The regulatory targets of the core EP transcription factors are those genomic regions bound by these factors in a coordinated manner in vivo (Boyer et al. 2005; Loh et al. 2006). The results presented above predict that these regions must have been assembled after the divergence of mammals from other vertebrates. To test this, we analyzed genome-wide those regions bound by the core EP transcription factors in ES cells (Marson et al. 2008) by examining whole-genome alignments of several vertebrate species, including mouse and chick (Frazer et al. 2004). We reasoned that if the core of the EP-GRN network (*Oct4*, *Nanog* and *Sox2*) is indeed a mammalian novelty, genomic regions that contain sites bound by these factors (most likely corresponding to cis-regulatory elements) will be poorly conserved between mouse and chick.

We examined the overall evolutionary conservation (see Materials and Methods) of 1688 non-coding genomic regions bound by, at least, *Oct4* and *Nanog* in ES cells (Marson et al. 2008). Of these regions, 11.55% are not conserved between mouse and any other species, over half (53.26%) are conserved only in rodents (mouse and rat), approximately a third (32.7%) are conserved in mammals (mouse, rat, human and dog), and just 2.49% are conserved in chick (Fig. 3A; Table S1).

This figure is extremely low compared with prior estimates of the conservation of non-coding regulatory elements between human, rodents and chick (Hillier et al. 2004). Nonetheless, to strengthen our conclusion we wished to compare the results with a similar dataset from a GRN known to control a biological process well-conserved between mouse and chick. For this, we choose the cis-regulatory circuitry underlying limb patterning, which is in many instances a prime example of the evolutionary conservation of developmental processes in vertebrates. Applying the same approach described above to the 200 top-scoring genomic regions bound by the limb development GRN transcription factor *Gli3* (Vokes et al. 2008), we found that 26% of these regions are conserved between mouse and chick (Fig. 3A; Table S1). It is also noteworthy that a higher proportion of the *Gli3*-bound regions are conserved in the other mammals analyzed (60.5%, versus 32.7% for *Oct4* and *Nanog*). Moreover, more of the mouse *Oct4* and *Nanog* bound regions are mouse or rodent specific (but see below). This comparison shows that regions bound by *Oct4* and *Nanog*, and therefore part of the ES pluripotency GRN,

are significantly less conserved than those bound by Gli3 as part of the limb development GRN (in all cases, $p < 0.0001$; two-tailed Fisher's exact test). We can therefore propose that the EP-GRN was newly assembled during the evolution of mammals, in part through the recruitment of pre-existing genes by means of the acquisition of core EP factor response elements.

However, there are at least two alternative explanations for the data presented above. One possibility is that the majority of regions bound by Oct4 and Nanog in the mouse genome are cis-regulatory elements controlling mouse-specific genes that have no chick orthologues. Therefore our analysis would only detect mammalian or mouse specific EP-GRN targets. This is not the case, since we found that those genes located in the vicinity of Oct4-Nanog bound regions have a higher proportion of 1-to-1 chick orthologues than the total mouse gene set (p -value < 0.0001 ; two-tailed Chi square test; data not shown). The second possibility is that the low conservation of Oct4-Nanog regions in the chick genome could simply be an artefact of their enrichment in mouse- or rodent-only conserved regions (see above; Fig. 3A). To test for this, we first counted the number of Oct4-Nanog regions conserved between mouse and human, and from this set we counted those regions that were also conserved with rat or with chick. We then calculated the relative proportions of each grouping (Table S2). Applying this same strategy to whole genome comparison, it has been shown that 23.29% of all non-coding regions conserved among mouse, human and rat are also conserved in chick (Prabhakar et al. 2006). We found that for Oct4-Nanog regions the figure was 8.02% and for Gli3 regions it was 32.3% (Fig. 3B; Table S2). Therefore the lack of conservation of Oct4-Nanog bound regions between mouse and chick is still significant when mouse and rodent-specific genomic regions are excluded from the analysis (p -value < 0.0001 ; two-tailed Fisher's exact test).

Conservation and evolution of the *Fgf4* and *Sox2* EP enhancers

To refine this scenario, we tried to identify specific cases where we could trace the appearance of a core EP-GRN regulated element. For this, we focused on elements that have been shown to possess cis-regulatory activity in vivo in response to core EP factors. The best characterized set of such regulatory elements corresponds to Sox2/Oct4 regulated genes in which adjacent HMG and POU binding motifs mediate high transcriptional activity (Remenyi et al. 2003). To date, functional HMG/POU cassettes have been characterized in detail from eight Sox2/Oct4 target genes, namely *Sox2* and *Oct4* themselves plus *Nanog*, *Lefty1*, *Fgf4*, *Fbxo15*, *Utf1* and *Dppa4* (Nakatate et al. 2006; Chakravarthy et al. 2008) and references therein).

As already shown above, *Oct4*, *Utf1* and *Dppa4* are not present in the chick (Figs S1; 2A, D), so the conservation of the sites for these genes can not be examined in an orthologous genomic context. In the case of *Nanog*, *Lefty1* and *Fbxo15*, the regions spanning the Sox2/Oct4 sites contain no regions of non-coding sequence conservation between mouse and chick. Finally, for both *Fgf4* and *Sox2* we found that the genomic region containing the HMG/POU cassette shows a partial but significant degree of conservation between mouse, other mammals and chick (Fig. 4A, B). This allowed us to unambiguously align the regions, using peaks of high conservation as

landmarks, and examine the conservation of the Sox2/Oct4 binding sites and surrounding sequences in detail.

The *Fgf4* HMG/POU cassette is located in the 3' UTR of the gene, and genomic fragments carrying these sites drive reporter expression in embryonic carcinoma (EC) cells, ES cells and the ICM of hatching blastocysts (4.5 dpc), and this expression is strictly dependant on the synergistic action of Sox2 and Oct4 (Yuan et al. 1995; Fraidenraich et al. 1998; Chakravarthy et al. 2008). Since the 3' UTR of the *Fgf4* gene is not described in many species, to examine the conservation of these sites and surrounding regions we aligned the 3' UTR of mouse *Fgf4* (2.3 kb) to 3 kb of genomic sequence immediately downstream from the stop codon of the coding region of *Fgf4* orthologues from eutherian mammals (rat, dog, cow and human), a marsupial (opossum), a monotreme (platypus), chick and zebrafish (Fig. 4A). We found various peaks of conservation distributed along the length of the mouse 3' UTR and noticed that the Sox2/Oct4 binding site (thin blue line in Fig. 4A) was located in a region conserved among eutherian mammals, absent from opossum and zebrafish, and weakly conserved in platypus and chick. In these two last cases, sequence conservation was sufficient to anchor an alignment that contained the HMG/POU cassette (Fig. 4C). The HMG half of the cassette is identical in all species except for a single change in the 5' end in the chick sequence (A instead of C). Intriguingly, this is one of the invariable residues in all Sox2/Oct4 composite sites identified to date in mouse (Chew et al. 2005). Regarding the POU site, both chick and platypus show nucleotide changes and a one base deletion. Based on studies of other Sox2/Oct4 elements, these changes would abolish Oct4 binding and enhancer activity (Chew et al. 2005; Kuroda et al. 2005; Rodda et al. 2005; Chakravarthy et al. 2008).

The Sox2 HMG/POU cassette localizes approximately 2.5 kb downstream of the stop codon in mouse (Tomioka et al. 2002). We therefore aligned the 4 kb downstream from the stop codon of the mouse gene to the equivalent regions from the same species used above. As for *Fgf4*, conservation was distributed along the length of the sequence and the Sox2/Oct binding site was located in a region highly conserved among all mammals, and poorly but still detectably conserved in chick (Fig. 4B). This region does not correspond to any of the multiple conserved elements previously described to act as enhancers in the early chick embryo (Uchikawa et al. 2003). Closer examination showed that unlike the *Fgf4* enhancer, the Sox2 enhancer of non-eutherian mammals contains a perfect POU site, and the HMG site contains only two changes; however, once again these affect the C residue in the first position, invariant in all other Sox2/Oct4 sites (Fig. 5D). The chick sequence is altered with respect to the mouse at multiple positions in both the HMG and the POU sites (Fig. 4D). Experimental evidence has shown that changes at many of these positions, in some cases combined with changes in downstream positions that are neither conserved between chick and mouse, severely reduce enhancer activity (Fig. S2) (Tomioka et al. 2002; Chakravarthy et al. 2008). This suggests that the chick sequence would be unable to bind Sox2 and Oct4, and therefore would not act as an enhancer element in response to these factors.

This analysis has thus identified two examples in which overall similarity allows us to unambiguously identify the chick genomic region that corresponds to mouse Sox2/Oct4 responsive elements active during pluripotency, but which lack critical sequences required for binding by Sox2 and Oct4. The number of residues that change from chick to mouse is small enough to suggest that changes to these sequences by point mutation and insertion led to the appearance of novel regulatory elements under the control of Sox2 and Oct4, as part of the emerging EP-GRN during mammalian evolution.

DISCUSSION

Pluripotency is a necessary and transient state in any multicellular organism that develops through an obligatory one-cell stage as part of its life cycle. However, our knowledge about the genetic control of pluripotency in the embryo is very much limited to the mouse. This is mainly due to the unique opportunity that embryonic stem (ES) cells offer us, as they represent a frozen state in development where the gene regulatory networks responsible for embryonic pluripotency (EP-GRNs) are maintained indefinitely, and the activation of differentiation programmes can be controlled experimentally by culture conditions (Niwa 2007; Silva and Smith 2008).

In this way, the main factors and signals involved in embryonic pluripotency of mouse ES cells and the blastocyst have been identified (Boiani and Scholer 2005). This knowledge has been used to analyze this process in other mammals such as human, pig or ruminants. Although some interesting differences have been found (Xu et al. 2002; Kuijk et al. 2008), this comparison has provided little insight into how deeply conserved is the EP-GRN, and how did it appear during evolution. Recent studies in non-mammalian vertebrates have begun to address this issue, mainly by studying orthologues of the core EP-GRN transcription factors (Canon et al. 2006; Morrison and Brickman 2006; Laval et al. 2007), but a clear understanding of the evolution of embryonic pluripotency is still missing. In order to fill this gap, we have undertaken a comparative genomic and gene expression analysis of the EP-GRN between mouse and chick, the closest vertebrate species where sufficient reagents and embryonic material are available to study the process in detail.

Early patterning in the chick embryo

The freshly laid chicken egg contains an embryo that has not started gastrulation and where two distinct concentric regions can be observed: the central area pelucida, from which the embryo proper will develop; and the exterior area opaca, that will form the majority of the extraembryonic membranes. Shortly after laying, the primary hypoblast will appear as a result of delamination of cells from the area pelucida. This arrangement is topologically similar to the mouse blastocyst, with the separation of embryonic and extraembryonic lineages and the

subsequent formation of the primitive endoderm, and in fact has been proposed to align more precisely the early developmental stages of vertebrate embryos (O'Farrell et al. 2004). Surprisingly, we found little evidence for comparable expression of orthologues of mouse EP-GRN genes in stage EGK-X chick embryos (Fig. 1). *Sox2*, *Fgf4* or *Nodal*, are not expressed until post-gastrulation stages, in stark contrast to their mouse orthologues that show expression in the ICM of the blastocyst (Niswander and Martin 1992; Avilion et al. 2003; Takaoka et al. 2006). A previous report has shown that a chick *Oct4* homologue is expressed ubiquitously in the epiblast of stage EGK-XI embryos (Lavial et al. 2007), suggesting a similar role in embryonic pluripotency to that of mouse. However, this gene is not a true homologue but a paralogue of mouse *Oct4* (Niwa et al. 2008). Similarly, we previously showed that chick *Nanog* was expressed only in primordial germ cells of the pre-gastrulation chick embryo (Canon et al. 2006).

This is an important observation, as by this stage (EGK-X) the chick embryo is already spatially patterned (Eyal-Giladi and Kochav 1976; Kochav et al. 1980; Callebaut 2005) and is composed by tens of thousands of cells, in comparison to the hundred(s) of the mouse blastocyst. The lack of expression of the orthologues of the key EP-GRN genes rules out them having a role in establishing the epiblast fate of the chick embryo prior to gastrulation as happens in mouse. It is true that we have not examined if these genes are expressed at earlier intrauterine stages in the chick, where they could be exerting a role in embryonic pluripotency and lineage specification. Such scenario would necessarily imply maternal expression of these genes, as zygotic genome activation occurs at EGK-X stage (Zagris et al. 1998; Elis et al. 2008), and the posterior asymmetric distribution of transcripts and/or proteins between cell populations. Certainly, maternal expression of the zebrafish *Pou5f1*-related gene *pou2* is required for early endodermal development (Reim and Brand 2006), and mouse *Oct4* is expressed maternally in the oocyte (Ovitt and Scholer 1998). Nevertheless, the role in pluripotency of the mouse gene is related to its zygotic expression, and the restricted expression of *Oct4* in the blastocyst is under the control of regulatory elements that are activated zygotically and not in the oocyte (Yeom et al. 1996). Taken together, the comparative expression analysis between pre-gastrulation mouse and chick embryos suggests that the EP-GRN was newly assembled in mammals, where expression of core transcription factors appears in novel domains in space and time during evolution.

Recruiting genes to the EP-GRN

Some of the genes that are part of the mouse EP-GRN are novel and specific to mammals and in many instances were generated by gene duplication, such as *Gdf3* (Canon et al. 2006) or *Rex1* (Kim et al. 2007). It is tempting to speculate that the appearance of novel genes as a result of gene duplication was a mayor event behind the emergence of the pluripotency regulatory network, as has been suggested in the case of for *Oct4* (Niwa et al. 2008). However, when we examined those genes adjacent to genomic regions bound by Oct4 and Nanog in ES

cells (Marson et al. 2008), and as such strong candidates to be direct downstream targets of these factors, we find that they are not enriched in mammalian specific genes. Furthermore, here we show that the duplication of *Oct4* occurred before the divergence of amniotes, indicating that all core pluripotency factors were present long before the appearance of mammals. This means that if the EP-GRN was newly assembled in mammals, as the expression data suggests, it must have occurred mainly through the recruitment of pre-existing genes.

The most obvious way in which genes can gain new expression domains is through the appearance of novel cis-regulatory elements in their vicinity. We find that the genomic regions bound by the core EP-GRN factors show very little conservation between mouse and chick, in accordance with them being new elements in the mammalian lineage. Further evidence for this comes from the fact that these regions are enriched in rodent- and mouse-specific sequences, what reflects their recent evolutionary appearance in the vertebrate lineage. On the contrary, this is not the case for *Gli3* bound regions part of the limb-specification GRN (Vokes et al. 2008), that are highly conserved in mammals.

We have been able to find two examples which confirm our predictions. For both *Fgf4* and *Sox2* we found that the sequences surrounding the *Sox2-Oct4* cassette, critical for their expression in the pre-implantation embryo and ES cells (Yuan et al. 1995; Fraidenraich et al. 1998; Tomioka et al. 2002), are sufficiently conserved between mammals and chick to allow the identification critical changes in residues that result in the appearance of adjacent functional HMG and POU binding sites. *More depending on the enhancer exp.* It could be argued that the nucleotide differences we observe in the chick correlate with changes in the binding specificities of the factors. However, recent work has shown that the chick *Oct4*-related gene *Pou2-r* is able to substitute the endogenous *Oct4* in mouse ES cells (Lavial et al. 2007), and the alignment of the DNA binding domains of these factors does not reveal amino acid substitutions in chick that would dramatically alter binding (not shown).

What is special about embryonic pluripotency in mammals?

In light of our results, we may ask why a new gene regulatory network in charge of maintaining embryonic pluripotency would have arisen in mammals. We believe that this may be due to the peculiarities of the early mammalian embryo and its nourishment requirements. In the majority of other vertebrate (and invertebrate) embryos, the egg is supplied with yolk that will feed the embryo through development. In these cases, the embryo grows by a series of quick cell divisions (O'Farrell et al. 2004), and patterning is initiated shortly after fertilization triggered by maternal and/or external factors (such as the sperm entry point in amphibians, or gravitation in the chick).

The situation is radically different in mammals, where eggs are devoid of yolk and embryo nourishment must be obtained from direct contact with the maternal uterine wall via the

placenta. This change has been accompanied by the loss of genes coding for vitellogenin proteins, stored in the yolk as nutritional reservoir (Brawand et al. 2008), and the appearance of a new cellular population, the trophoblast, that will give rise to the majority of the placenta. In this regard, we have also found striking differences in expression when studying orthologues of trophoblast-specific genes in the early chick embryo (Pernaute et al., in preparation). Initial cell divisions of the embryo, as observed in mouse, are extremely slow (O'Farrell et al. 2004), and until implantation there is hardly any growth. As with regards to embryonic patterning, it now seems clear that it is not initiated until peri-implantation stages, long after fertilization (Rossant and Tam 2009). Interestingly enough, another characteristic of the mouse embryo is that the initiation of patterning begins due to intrinsic and zygotic mechanisms: no external trigger or maternal signal is needed to establish the major body axes. This is most clearly viewed in tetraploid aggregation chimeras, where nearly the whole embryo is ES cell derived (Eakin and Hadjantonakis 2006), and these can not possess or retain any spatial patterning information after prolonged growth and passages under tissue culture conditions.

It is under these conditions, where a delay in pattern formation and an uncoupling from external cues occurs, that it is necessary to maintain embryonic cells in an undetermined and quiescent state for a prolonged period. The appearance of the EP-GRN would permit such state, where the core factors (Oct4-Sox2-Nanog) maintain genes involved in early differentiation and specification silent (Boyer et al. 2005; Loh et al. 2006) but these are at the same time ready to be expressed in a controlled spatial and temporal fashion. Such situation is attained by the unique epigenetic status of ES cells, where these genes are marked by both activation and repressive marks in so called bivalent domains (Bernstein et al. 2006), and RNA polymerase II is assembled at their promoters in a poised state (Stock et al. 2007).

The expression of *Nanog* and *Pou5f1* paralogues in PGC of both mouse and chick embryos (Canon et al. 2006; Laval et al. 2007; Rodriguez-Leon et al. 2008) can shed some light on the origin of the EP-GRN. PGC resemble ES cells in many aspects (Zwaka and Thomson 2005; Seydoux and Braun 2006), and a possible scenario would be that *Nanog* and *Pou5*-like genes would act in PGC to prevent their differentiation and maintain their germline potential until differentiation of the mature germ cells. The necessity of both these factors for germ cell development in mouse has been shown (Kehler et al. 2004; Chambers et al. 2007). Therefore, pluripotency maintenance in PGC would represent an evolutionary conserved role in vertebrates, that was later co-opted and the network expanded for the same function but in early embryonic cells of mammals.

MATERIALS AND METHODS

Embryos and in situ hybridization

Embryos were staged according to Eyal-Giladi and Kochav (Eyal-Giladi and Kochav 1976) for pre-gastrulation stages (EGK series) and to Hamburger and Hamilton (Hamburger and Hamilton 1992) for post-gastrulation stages (HH series).

Whole mount in situ hybridization of chick embryos was carried out as described (Ariza-McNaughton and Krumlauf 2002; Acloque et al. 2008). The chick *Pou2-r* clone was cloned by RT-PCR from genital ridges of stage HH22 embryos, the *Fgf4* and *Nodal* cDNAs a kind gift of Dr. Paola Bovolenta, and the *Sox2* (clone ChEST878b12, acc. n° BU282995), *FoxD3* (clone ChEST68l3, acc. n° BU128393) and *Phc1* (clone ChEST49d22, acc. n° BU219008) cDNAs were obtained from the MRC Geneservices (Boardman et al. 2002). In all cases, embryos of different stages were processed in parallel and sense probes used as negative controls.

Sequence analysis

Chromosomal position and genomic organization of chick and other vertebrates genes and genomic sequences from the *Fgf4* and *Sox2* locus from different species were obtained from the latest release of the Ensembl genome browser (www.ensembl.org). Synteny analysis was aided by the use of the Genomicus browser, developed by the Dyogen Lab in Paris (www.dyogen.ens.fr). Orthology maps (Fig. S1A) were constructed by examining the presence of 1-to-1 orthologues in different species (opossum, chick, *Xenopus* and zebrafish) of all genes in a 10 Mb region surrounding mouse *Oct4* and *Nanog*. Mouse genes with no orthologues in the species examined were not used and in the case of multiple matches to mouse genes in other species were reduced to only one.

Genomic regions bound by Oct4 and Nanog in mouse ES cells were obtained from the ChIP-seq data from Marson et al (2008). We have included in the analysis regions bound together by four (Oct4, Nanog, Sox2 and Tcf3; 814 regions), three (Oct4, Nanog and Sox2; 638 regions) or two (Nanog and Oct4; 236 regions) core EP-GRN transcription factors, located at 8 kb or less from known genes. We obtained equal results if each set was analyzed independently. Genomic regions bound by Gli3 in the mouse limb were obtained from the ChIP-seq data from Volkes et al. (2008). We used for the analysis the 200 top scoring regions as described in the paper. The coordinates for both data sets were converted from the mouse mm8 (February 2006) to the mm7 (August 2005) assembly, visualized using the Vista Tracks in the UCSC Browser (Frazer et al. 2004), and compared to rat, human, dog and chick. Each genomic region was individually analyzed and scored as conserved if it contained at least one segment of more than 100 bp and over 70% similarity in intergenic, intronic or non-translated regions. Conserved regions overlapping coding exons were not included. Regions were scored as mammalian specific if they were conserved between mouse and at least one of the two non-rodent species analyzed (human and dog).

Multi-species genomic alignments were performed using the Vista tools (Frazer et al. 2004), and sequences surrounding the Sox2/Oct4 sites were aligned using Clustal X (Thompson et al. 1997) and edited with BioEdit (Hall 1999).

ACKNOWLEDGEMENTS

We wish to thank Robb Krumlauf for reporter constructs, Paola Bovolenta for in situ probes, Tristan Rodriguez and Miguel Torres for insightful comments and suggestions, Eva Alonso and Cristina Arias for support and discussion, and Simon Bartlett for reading and correcting the manuscript. This work was supported by grants from the Spanish Government (BFU2008-00838 and CONSOLIDER-25120), the Regional Government of Madrid (CAM S-SAL-0190-2006), and by the ProCNIC Foundation.

REFERENCES

- Acloque H, Wilkinson DG, Nieto MA (2008) In situ hybridization analysis of chick embryos in whole-mount and tissue sections. *Methods Cell Biol* 87: 169-185.
- Ariza-McNaughton L, Krumlauf R (2002) Non-radioactive in situ hybridization: simplified procedures for use in whole-mounts of mouse and chick embryos. *Int Rev Neurobiol* 47: 239-250.
- Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N et al. (2003) Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev* 17(1): 126-140.
- Bachvarova RF, Masi T, Drum M, Parker N, Mason K et al. (2004) Gene expression in the axolotl germ line: *Axdazl*, *Axvh*, *Axoct-4*, and *Axkit*. *Dev Dyn* 231(4): 871-880.
- Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ et al. (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125(2): 315-326.
- Boardman PE, Sanz-Ezquerro J, Overton IM, Burt DW, Bosch E et al. (2002) A comprehensive collection of chicken cDNAs. *Curr Biol* 12(22): 1965-1969.
- Boiani M, Scholer HR (2005) Regulatory networks in embryo-derived pluripotent stem cells. *Nat Rev Mol Cell Biol* 6(11): 872-884.
- Bortvin A, Eggan K, Skaletsky H, Akutsu H, Berry DL et al. (2003) Incomplete reactivation of Oct4-related genes in mouse embryos cloned from somatic nuclei. *Development* 130(8): 1673-1680.
- Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS et al. (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122(6): 947-956.
- Brawand D, Wahli W, Kaessmann H (2008) Loss of egg yolk genes in mammals and the origin of lactation and placentation. *PLoS Biol* 6(3): e63.
- Burgess S, Reim G, Chen W, Hopkins N, Brand M (2002) The zebrafish *spiel-ohne-grenzen* (*spg*) gene encodes the POU domain protein Pou2 related to mammalian Oct4 and is essential for formation of the midbrain and hindbrain, and for pre-gastrula morphogenesis. *Development* 129(4): 905-916.
- Callebaut M (2005) Origin, fate, and function of the components of the avian germ disc region and early blastoderm: role of ooplasmic determinants. *Dev Dyn* 233(4): 1194-1216.
- Canon S, Herranz C, Manzanares M (2006) Germ cell restricted expression of chick Nanog. *Dev Dyn* 235(10): 2889-2894.
- Clark AT, Rodriguez RT, Bodnar MS, Abeyta MJ, Cedars MI et al. (2004) Human STELLAR, NANOG, and GDF3 genes are expressed in pluripotent cells and map to chromosome 12p13, a hotspot for teratocarcinoma. *Stem Cells* 22(2): 169-179.
- Chakravarthy H, Boer B, Desler M, Mallanna SK, McKeithan TW et al. (2008) Identification of DPPA4 and other genes as putative Sox2:Oct-3/4 target genes using a combination of in silico analysis and transcription-based assays. *J Cell Physiol* 216(3): 651-662.
- Chambers I, Silva J, Colby D, Nichols J, Nijmeijer B et al. (2007) Nanog safeguards pluripotency and mediates germline development. *Nature* 450(7173): 1230-1234.
- Chapman SC, Schubert FR, Schoenwolf GC, Lumsden A (2002) Analysis of spatial and temporal gene expression patterns in blastula and gastrula stage chick embryos. *Dev Biol* 245(1): 187-199.
- Chen X, Xu H, Yuan P, Fang F, Huss M et al. (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133(6): 1106-1117.
- Chew JL, Loh YH, Zhang W, Chen X, Tam WL et al. (2005) Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol Cell Biol* 25(14): 6031-6046.
- Deakin JE, Papenfuss AT, Belov K, Cross JG, Coggill P et al. (2006) Evolution and comparative analysis of the MHC Class III inflammatory region. *BMC Genomics* 7: 281.
- Eakin GS, Hadjantonakis AK (2006) Production of chimeras by aggregation of embryonic stem cells with diploid or tetraploid mouse embryos. *Nat Protoc* 1(3): 1145-1153.
- Elis S, Batellier F, Couty I, Balergue S, Martin-Magniette ML et al. (2008) Search for the genes involved in oocyte maturation and early embryo development in the hen. *BMC Genomics* 9: 110.
- Eyal-Giladi H, Kochav S (1976) From cleavage to primitive streak formation: a complementary normal table and a new look at the first stages of the development of the chick. I. General morphology. *Dev Biol* 49(2): 321-337.
- Fraidenraich D, Lang R, Basilico C (1998) Distinct regulatory elements govern Fgf4 gene expression in the mouse blastocyst, myotomes, and developing limb. *Dev Biol* 204(1): 197-209.

- Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I (2004) VISTA: computational tools for comparative genomics. *Nucleic Acids Res* 32(Web Server issue): W273-279.
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* 41: 95-98.
- Hamburger V, Hamilton HL (1992) A series of normal stages in the development of the chick embryo. 1951. *Dev Dyn* 195(4): 231-272.
- Hanna LA, Foreman RK, Tarasenko IA, Kessler DS, Labosky PA (2002) Requirement for Foxd3 in maintaining pluripotent cells of the early mouse embryo. *Genes Dev* 16(20): 2650-2661.
- Hillier LW, Miller W, Birney E, Warren W, Hardison RC et al. (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 432(7018): 695-716.
- Houbaviy HB, Murray MF, Sharp PA (2003) Embryonic stem cell-specific MicroRNAs. *Dev Cell* 5(2): 351-358.
- Houbaviy HB, Dennis L, Jaenisch R, Sharp PA (2005) Characterization of a highly variable eutherian microRNA gene. *Rna* 11(8): 1245-1257.
- Hughes AL, Friedman R (2008) Genome size reduction in the chicken has involved massive loss of ancestral protein-coding genes. *Mol Biol Evol* 25(12): 2681-2688.
- Inoue N, Hess KD, Moreadith RW, Richardson LL, Handel MA et al. (1999) New gene family defined by MORC, a nuclear protein required for mouse spermatogenesis. *Hum Mol Genet* 8(7): 1201-1207.
- Kehler J, Tolkunova E, Koschorz B, Pesce M, Gentile L et al. (2004) Oct4 is required for primordial germ cell survival. *EMBO Rep* 5(11): 1078-1083.
- Kim J, Chu J, Shen X, Wang J, Orkin SH (2008) An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* 132(6): 1049-1061.
- Kim JD, Faulk C, Kim J (2007) Retroposition and evolution of the DNA-binding motifs of YY1, YY2 and REX1. *Nucleic Acids Res* 35(10): 3442-3452.
- Kochav S, Ginsburg M, Eyal-Giladi H (1980) From cleavage to primitive streak formation: a complementary normal table and a new look at the first stages of the development of the chick. II. Microscopic anatomy and cell population dynamics. *Dev Biol* 79(2): 296-308.
- Kos R, Reedy MV, Johnson RL, Erickson CA (2001) The winged-helix transcription factor FoxD3 is important for establishing the neural crest lineage and repressing melanogenesis in avian embryos. *Development* 128(8): 1467-1479.
- Kuijk EW, Du Puy L, Van Tol HT, Oei CH, Haagsman HP et al. (2008) Differences in early lineage segregation between mammals. *Dev Dyn* 237(4): 918-927.
- Kuntz S, Kieffer E, Bianchetti L, Lamoureux N, Fuhrmann G et al. (2008) Tex19, a mammalian-specific protein with a restricted expression in pluripotent stem cells and germ line. *Stem Cells* 26(3): 734-744.
- Kuroda T, Tada M, Kubota H, Kimura H, Hatano SY et al. (2005) Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression. *Mol Cell Biol* 25(6): 2475-2485.
- Lavial F, Acloque H, Bertocchini F, Macleod DJ, Boast S et al. (2007) The Oct4 homologue PouV and Nanog regulate pluripotency in chicken embryonic stem cells. *Development* 134(19): 3549-3563.
- Loh YH, Wu Q, Chew JL, Vega VB, Zhang W et al. (2006) The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet* 38(4): 431-440.
- Maldonado-Saldivia J, van den Bergen J, Krouskos M, Gilchrist M, Lee C et al. (2007) Dppa2 and Dppa4 are closely linked SAP motif genes restricted to pluripotent cells and the germ line. *Stem Cells* 25(1): 19-28.
- Marson A, Levine SS, Cole MF, Frampton GM, Brambrink T et al. (2008) Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* 134(3): 521-533.
- Masaki H, Nishida T, Kitajima S, Asahina K, Teraoka H (2007) Developmental pluripotency-associated 4 (DPPA4) localized in active chromatin inhibits mouse embryonic stem cell differentiation into a primitive ectoderm lineage. *J Biol Chem* 282(45): 33034-33042.
- Morrison GM, Brickman JM (2006) Conserved roles for Oct4 homologues in maintaining multipotency during early vertebrate development. *Development* 133(10): 2011-2022.
- Nakatake Y, Fukui N, Iwamatsu Y, Masui S, Takahashi K et al. (2006) Klf4 cooperates with Oct3/4 and Sox2 to activate the Lefty1 core promoter in embryonic stem cells. *Mol Cell Biol* 26(20): 7772-7782.

- Niswander L, Martin GR (1992) Fgf-4 expression during gastrulation, myogenesis, limb and tooth development in the mouse. *Development* 114(3): 755-768.
- Niwa H (2007) How is pluripotency determined and maintained? *Development* 134(4): 635-646.
- Niwa H, Sekita Y, Tsend-Ayush E, Grutzner F (2008) Platypus Pou5f1 reveals the first steps in the evolution of trophectoderm differentiation and pluripotency in mammals. *Evol Dev* 10(6): 671-682.
- O'Farrell PH, Stumpff J, Su TT (2004) Embryonic cleavage cycles: how is a mouse like a fly? *Curr Biol* 14(1): R35-45.
- Okuda A, Fukushima A, Nishimoto M, Orimo A, Yamagishi T et al. (1998) UTF1, a novel transcriptional coactivator expressed in pluripotent embryonic stem cells and extra-embryonic cells. *Embo J* 17(7): 2019-2032.
- Ovitt CE, Scholer HR (1998) The molecular biology of Oct-4 in the early mouse embryo. *Mol Hum Reprod* 4(11): 1021-1031.
- Pierre A, Gautier M, Callebaut I, Bontoux M, Jeanpierre E et al. (2007) Atypical structure and phylogenomic evolution of the new eutherian oocyte- and embryo-expressed KHDC1/DPPA5/ECAT1/OOEP gene family. *Genomics* 90(5): 583-594.
- Popova NV, Morris RJ (2004) Genetic regulation of mouse stem cells: identification of two keratinocyte stem cell regulatory loci. *Curr Top Microbiol Immunol* 280: 111-137.
- Prabhakar S, Poulin F, Shoukry M, Afzal V, Rubin EM et al. (2006) Close sequence comparisons are sufficient to identify human cis-regulatory elements. *Genome Res* 16(7): 855-863.
- Reim G, Brand M (2006) Maternal control of vertebrate dorsoventral axis formation and epiboly by the POU domain protein Spg/Pou2/Oct4. *Development* 133(14): 2757-2770.
- Remenyi A, Lins K, Nissen LJ, Reinbold R, Scholer HR et al. (2003) Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers. *Genes Dev* 17(16): 2048-2059.
- Rex M, Orme A, Uwanogho D, Tointon K, Wigmore PM et al. (1997) Dynamic expression of chicken Sox2 and Sox3 genes in ectoderm induced to form neural tissue. *Dev Dyn* 209(3): 323-332.
- Rodda DJ, Chew JL, Lim LH, Loh YH, Wang B et al. (2005) Transcriptional regulation of nanog by OCT4 and SOX2. *J Biol Chem* 280(26): 24731-24737.
- Rodriguez-Leon J, Rodriguez Esteban C, Marti M, Santiago-Josefat B, Dubova I et al. (2008) Pitx2 regulates gonad morphogenesis. *Proc Natl Acad Sci U S A* 105(32): 11242-11247.
- Rossant J, Tam PP (2009) Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. *Development* 136(5): 701-713.
- Sayers EW, Barrett T, Benson DA, Bryant SH, Canese K et al. (2008) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res*.
- Seydoux G, Braun RE (2006) Pathway to totipotency: lessons from germ cells. *Cell* 127(5): 891-904.
- Shamim H, Mason I (1999) Expression of Fgf4 during early development of the chick embryo. *Mech Dev* 85(1-2): 189-192.
- Silva J, Smith A (2008) Capturing pluripotency. *Cell* 132(4): 532-536.
- Stock JK, Giadrossi S, Casanova M, Brookes E, Vidal M et al. (2007) Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. *Nat Cell Biol* 9(12): 1428-1435.
- Takaoka K, Yamamoto M, Shiratori H, Meno C, Rossant J et al. (2006) The mouse embryo autonomously acquires anterior-posterior polarity at implantation. *Dev Cell* 10(4): 451-459.
- Tanaka TS, Lopez de Silanes I, Sharova LV, Akutsu H, Yoshikawa T et al. (2006) Esg1, expressed exclusively in preimplantation embryos, germline, and embryonic stem cells, is a putative RNA-binding protein with broad RNA targets. *Dev Growth Differ* 17(6): 381-390.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25(24): 4876-4882.
- Tomioka M, Nishimoto M, Miyagi S, Katayanagi T, Fukui N et al. (2002) Identification of Sox-2 regulatory region which is under the control of Oct-3/4-Sox-2 complex. *Nucleic Acids Res* 30(14): 3202-3213.

- Tomotsune D, Shirai M, Takihara Y, Shimada K (2000) Regulation of Hoxb3 expression in the hindbrain and pharyngeal arches by rae28, a member of the mammalian Polycomb group of genes. *Mech Dev* 98(1-2): 165-169.
- Uchikawa M, Ishida Y, Takemoto T, Kamachi Y, Kondoh H (2003) Functional analysis of chicken Sox2 enhancers highlights an array of diverse regulatory elements that are conserved in mammals. *Dev Cell* 4(4): 509-519.
- Vokes SA, Ji H, Wong WH, McMahon AP (2008) A genome-scale analysis of the cis-regulatory circuitry underlying sonic hedgehog-mediated patterning of the mammalian limb. *Genes Dev* 22(19): 2651-2663.
- Xu RH, Chen X, Li DS, Li R, Addicks GC et al. (2002) BMP4 initiates human embryonic stem cell differentiation to trophoblast. *Nat Biotechnol* 20(12): 1261-1264.
- Yamanaka Y, Ralston A, Stephenson RO, Rossant J (2006) Cell and molecular regulation of the mouse blastocyst. *Dev Dyn* 235(9): 2301-2314.
- Yeom YI, Fuhrmann G, Ovitt CE, Brehm A, Ohbo K et al. (1996) Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. *Development* 122(3): 881-894.
- Yuan H, Corbi N, Basilico C, Dailey L (1995) Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. *Genes Dev* 9(21): 2635-2645.
- Zagris N, Kalantzis K, Guialis A (1998) Activation of embryonic genome in chick. *Zygote* 6(3): 227-231.
- Zwaka TP, Thomson JA (2005) A germ cell origin of embryonic stem cells? *Development* 132(2): 227-233.

FIGURE LEGENDS

Figure 1. Chick orthologues of mouse EP-GRN genes are not expressed in early pre-gastrulation embryos. *Pou2-r* is expressed in the earliest EGK-X stage embryos and shortly after is strongly upregulated in the area pelucida. *Sox2* is first expressed in the neural plate at HH5, *FoxD3* in the node at HH3, *Phc1* in the neural tube at later stages (HH9), *Fgf4* in the primitive streak at HH4, and *Nodal* in the early primitive streak and later restricted in the node (HH6). All of the mouse orthologues of these genes are expressed in the epiblast of the blastocyst (except for *Phc1*, whose expression in the early mouse embryo has not been reported) and in ES cells. See main text for details and references.

Figure 2. Lack of EP-GRN orthologues in the chick genome. Mouse *Utf1* (A), *Tex19.1* (B), *Dppa5* (C) and the tandem duplicates *Dppa2* and *Dppa4* (D) lie in regions of conserved synteny between mouse and the chick, but are absent from the later. The mouse (*Mm*) and chick (*Gg*) chromosomes (chr.) where the genes are located is indicated. The diagrams are not shown to scale.

Figure 3. Genomic regions bound by Oct4 and Nanog are poorly conserved in chick. (A) A high proportion of Oct4-Nanog bound regions (black bars) are specific to mouse or conserved only in rodents, compared to regions bound by Gli3 (white bars) as part of the limb development GRN. These last are also enriched in regions conserved between mammals. Very few of the Oct4-Nanog bound regions are conserved with chick, as compared to those bound by Gli3 (2.49 vs. 26%). (B) The set of Oct4-Nanog bound regions has an equal proportion of mammalian conserved non-coding elements compared to the Gli3 set, but a much lower amount of regions conserved up to chick. mh, number of genomic regions conserved between mouse and human; mhr, regions conserved between mouse and humans that are also conserved in rat; mhc, regions conserved between mouse and humans that are also conserved in chick. All differences between the Oct4-Nanog and the Gli3 data sets (except for the mhr/mh comparison) are highly significant (p value < 0.0001).

Figure 4. Conservation of the *Fgf4* and *Sox2* EP enhancers. (A, B) Vista plots of the 3'UTR from mouse *Fgf4* (A) and 4 kb downstream of the mouse *Sox2* stop codon (B) compared to different vertebrates. Regions of 100 bp or more that show 60% or higher sequence identity are coloured in pink, and the *Sox2*/*Oct4* binding site (BS) in blue. The zebrafish *Sox2* sequence does not any conservation with mouse and is not shown. (C, D) Sequence comparison of the regions surrounding the *Sox2*/*Oct4* BS from *Fgf4* (C) and *Sox2* (D). The HMG and POU sites are indicated below the sequence and with red background. Residues identical in at least five species are shown with a black background. Dashes indicate gaps introduced to maximize the

alignment. The opossum and zebrafish *Fgf4* sequences do not align in this region are not shown.

SUPPORTING INFORMATION

Figure S1. The genomic region containing *Oct4* is lost in avians. (A) Orthology maps show that the chick genome lacks orthologues for many genes that map in the vicinity of mouse (m) or opossum (o) *Oct4*, but are present in *Xenopus* (x) or zebrafish (z). By the contrary, genes surrounding mouse *Nanog* show a distribution of orthologues in different vertebrates that relates to their evolutionary distance from mouse. The position of mouse *Oct4* and *Nanog* is indicated by a green line. (B) Chick *Pou2-r* is located in a region of conserved synteny with that of other *Pou5f1*-related genes from non-mammalian vertebrates, but different from mouse *Pou5f1*. The opossum has two *Pou5f1*-like genes, one in conserved synteny with chick *Pou2-r*, and another with mouse *Oct4*, indicating that the mouse and chick genes are not orthologues but paralogues. In the case of the anole lizard, an orthologue of *Oct4* is present in conserved synteny with mouse, but not an orthologue of *Pou2-r*, surely due to lack of sequence coverage. The most parsimonious explanation for this distribution is that duplication on an ancestral *Pou5* gene occurred before the appearance of amniotes and subsequent losses occurred in avians (*Oct4*) and eutherian mammals (*Pou2-r*). This view is confirmed by phylogenetic analysis (not shown).

Figure S2. Alignment of the Sox2/Oct4 binding site and 3' adjacent sequences from mouse Sox2 with chick and mutant versions that significantly affect enhancer activity. All mutations were tested by Tomioka et al. (2002), except for the (Sox2)14G mutation that was tested by Chakravarthy et al. (2008). Dashes indicate gaps introduced to maximize the alignment and dots residues identical to the mouse sequence.

Tables S1 and S2.

Evolutionary conservation of genomic regions bound by Oct4 and Nanog in mouse ES cells, and by Gli3 in the mouse limb.

Figure 1

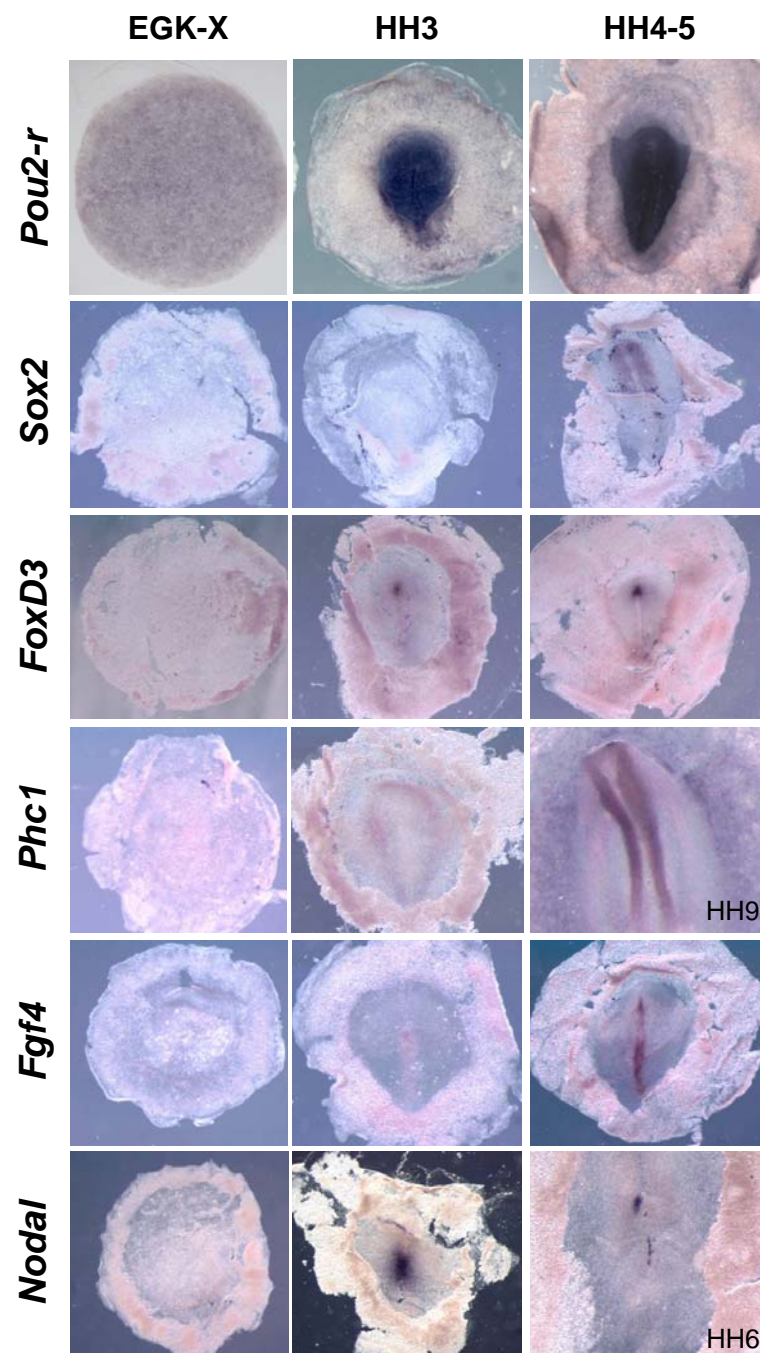


Figure 2

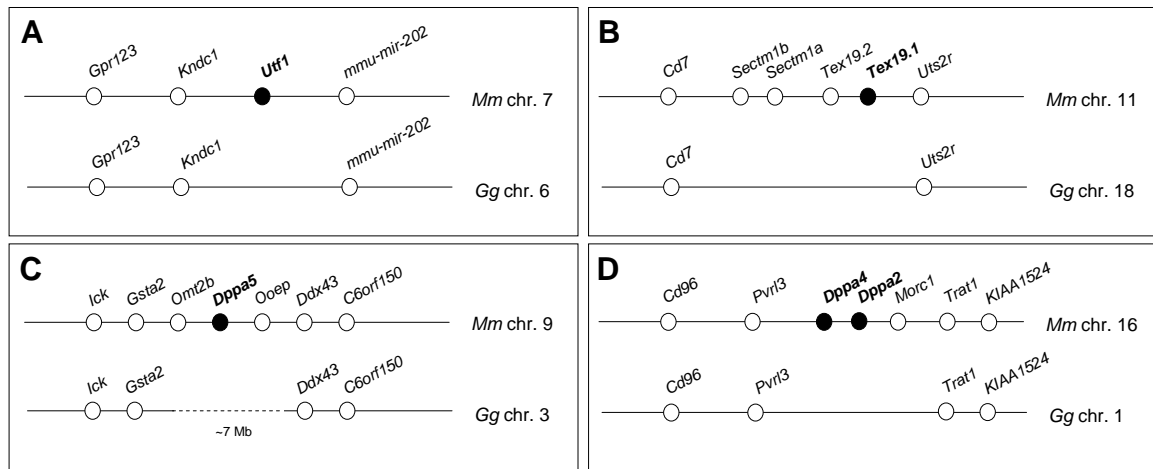


Figure 3

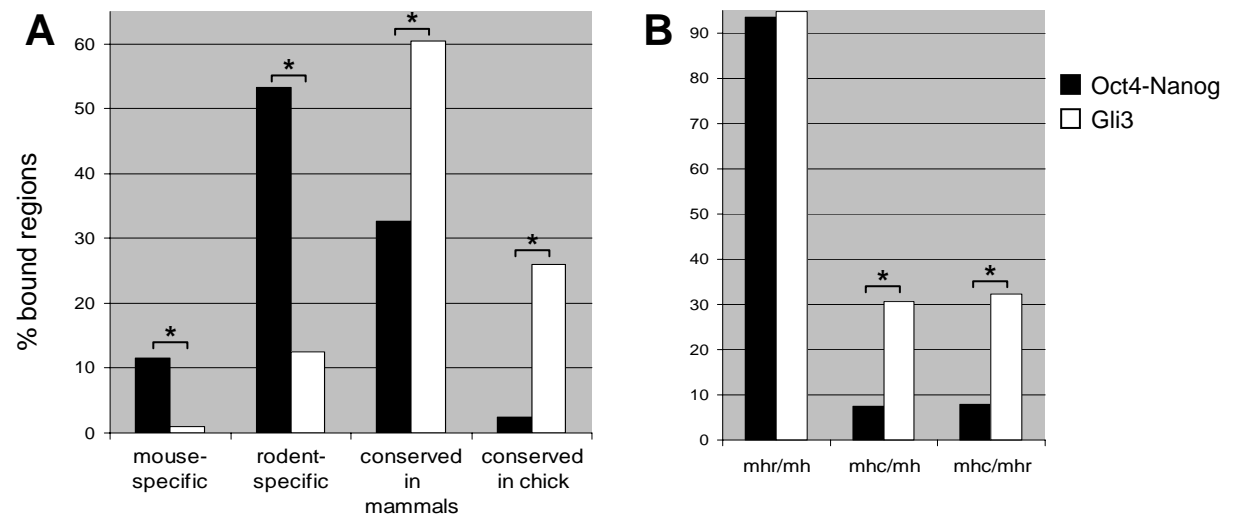


Figure 4

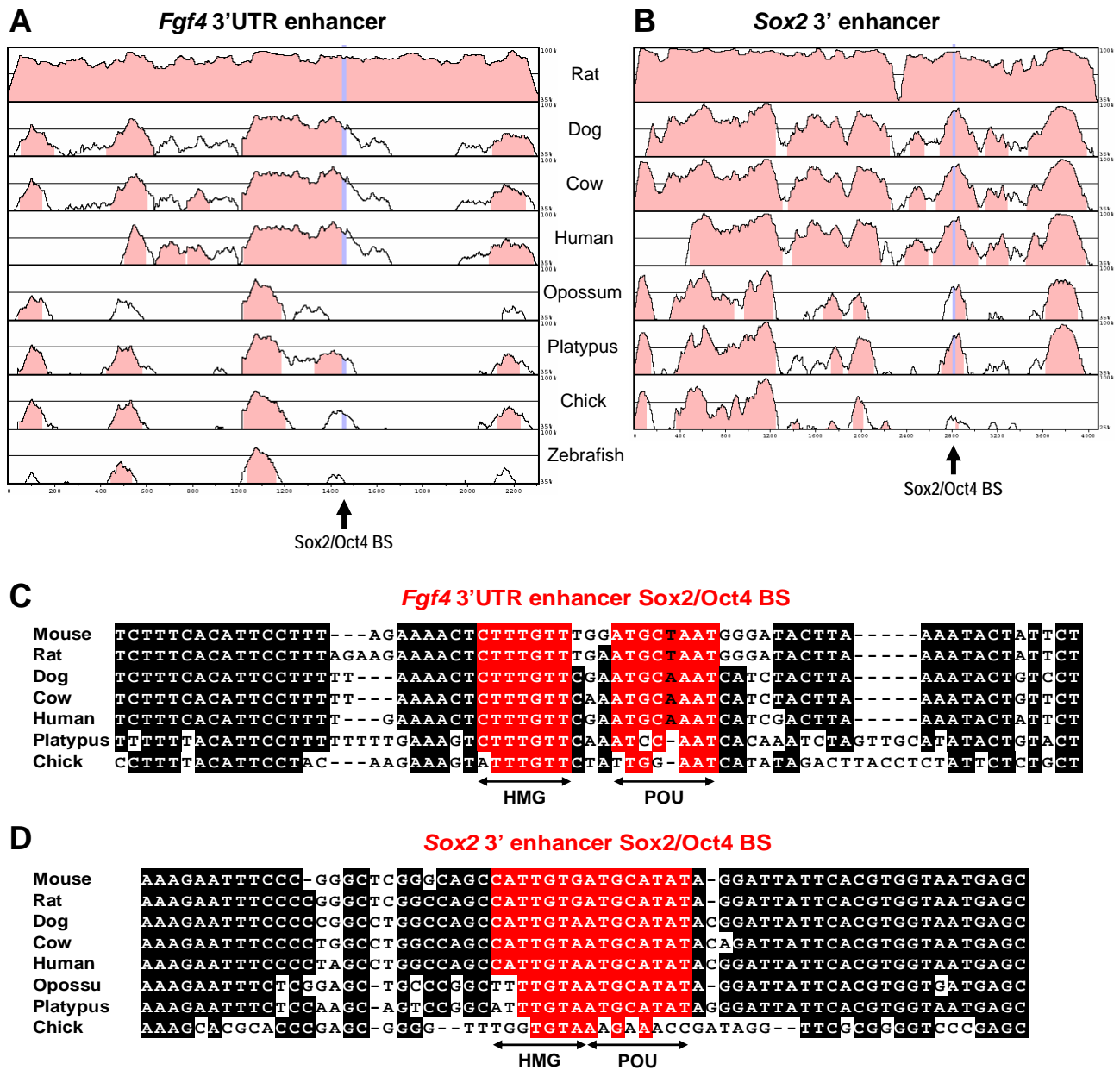
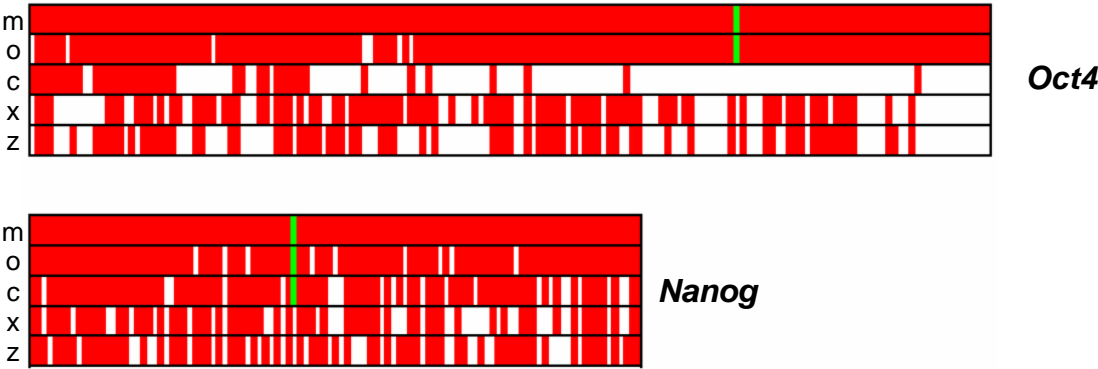
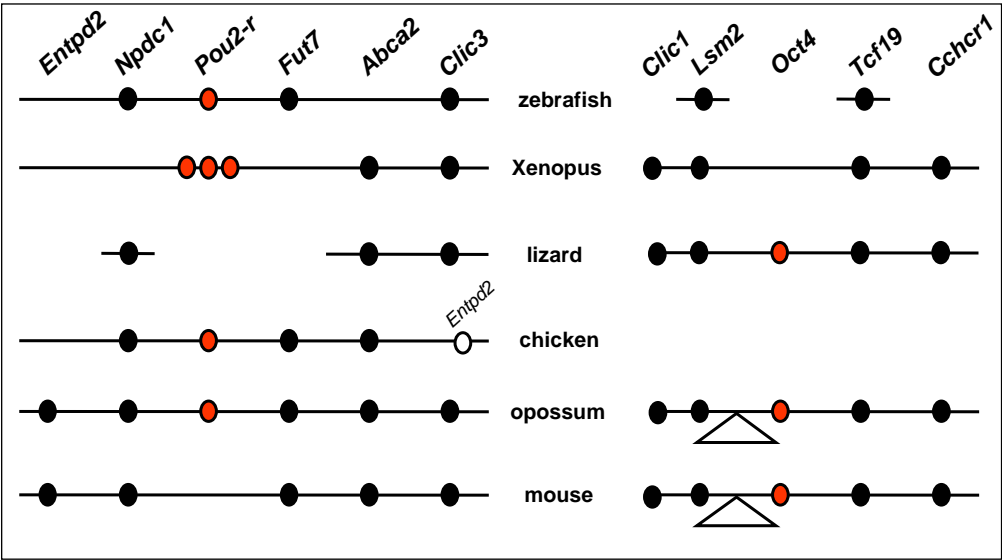


Figure S1

A



B



Mouse	CATTGTGATGCATATA	A-GGATTATT
Chick	TGG . . . A . A . A . ACC	GATAGG - - TT
<i>tpm16</i>	ACG
<i>tpm18</i> TCG
<i>muty</i> TAC CGGCGG
<i>mutδ</i> GCG CGGCGG
<i>(Sox2)14G</i> G

Figure S2. Alignment of the Sox2/Oct4 binding site and 3' adjacent sequences from mouse *Sox2* with chick and mutant versions that significantly affect enhancer activity. All mutations were tested by Tomioka et al. (2002), except for the *(Sox2)14G* mutation that was tested by Chakravarthy et al. (2008). Dashes indicate gaps introduced to maximize the alignment and dots residues identical to the mouse sequence.

Supporting Tables

Evolutionary conservation of genomic regions bound by Oct4 and Nanog in mouse ES cells, and by Gli3 in the mouse limb.

Table S1

	Oct4-Nanog		Gli3	
	n	%	n	%
mouse specific	195	11.55	2	1.00
rodent specific	899	53.26	25	12.50
mammalian specific	552	32.70	121	60.50
conserved in chick	42	2.49	52	26.00
Total	1688	100	200	100

Table S2

	Oct4-Nanog	Gli3
conserved mouse-human (mh)	560	170
conserved mouse-human-rat (mhr)	524	161
% conserved mh that also mhr	93.57	94.71
conserved mouse-human-chick (mhc)	42	52
% conserved mh that also mhc	7.5	30.59
% conserved mhr that also mhc	8.02	32.30

Detrás de los resultados presentados en esta tesis se encuentra el trabajo no sólo mío sino de toda la gente que ha colaborado en los dos proyectos principales en los que he trabajado durante estos cinco años y que tienen nombre de genes: *Dicer* y *Eomes*. A ellos, así como a todas aquellas personas cuyo apoyo científico y personal durante ese tiempo ha hecho que llegase hasta aquí van dedicadas estas últimas líneas.

Hace casi cinco años comencé la tesis cargada de ilusión y convencida de que el laboratorio era lo mío. Pasado este tiempo sigo con las mismas ganas con las que empecé o más, y buena parte de la culpa la tiene mi director de tesis: Miguel Manzanares. Muchas gracias Miguel por confiar en mí para formar parte de tu grupo, ha sido un placer tenerte como director no sólo por todo lo que me has aportado científicamente y todo lo que he aprendido, sino también por la experiencia personal que me llevo de estos años. Gracias por tu apoyo incondicional, por tu entusiasmo, por ser siempre tan positivo y por esforzarte en hacer que la gente que trabaja contigo disfrute de la ciencia como lo haces tú. Sabes que yo seguiría en este laboratorio por siempre jamás y espero que aunque no lo haga siga siempre contando de una manera u otra con tu apoyo.

Aunque no entraba entre mis planes el hacer la tesis en Londres al final resulta que la mayor parte del trabajo que presento en este libro lo he hecho allí bajo la supervisión de Tristán Rodríguez, que se ha convertido en mi co-director en esencia. Muchas gracias Tristán por acogerme en tu laboratorio durante casi un año y esforzarte en hacerme sentir una más desde el primer día. Muchas gracias por enseñarme de primera mano el trabajo con embriones que jamás pensé que sería capaz de diseccionar, y por ofrecerme analizar esos mutantes para *Dicer* que teníais por allí, nunca me imaginé que acabaría escribiendo una tesis sobre ellos. Muchas gracias por tu apoyo en todo el tiempo que estuve allí y por haber hecho que una y otra vez quiera volver.

Hay dos personas que han tomado parte directa en los resultados que presento en esta tesis y se merecen un agradecimiento especial porque el proyecto de *Dicer* no estaría saliendo adelante sin ellos: Aida y Thom, una parte importante de este trabajo es vuestra. Aida me dedicó todo su tiempo cuando estuve en Londres, y sigue dedicándomelo aun cuando me he ido ya que se quedó al cargo de todo lo que no pude terminar y que constituye parte de esta tesis. Muchas gracias Aida por ejercer de maestra, compañera, amiga y madre cuando estaba lejos de mi gente. Muchas gracias por enseñarme a "echarle piernas", porque todo se puede cuando hay ganas y merece la pena intentarlo. Gracias por ser como eres, única.

Muchas gracias a Thom sin el cual esta tesis no sería lo que es por muchos motivos. En primer lugar porque su tesis constituye la otra mitad de la historia de *Dicer*, y se hace difícil separar su trabajo del mío, por lo que parte de los datos que presento son suyos. Muchísimas gracias Thom por tu ayuda incondicional desde el día que me incorporé a trabajar codo con codo contigo. Muchas gracias por todo lo que has hecho y sigues haciendo porque este proyecto siga adelante y porque yo acabe esta tesis. Muchas gracias por tu infinita paciencia, apoyo y ayuda cuando me puse por objetivo acabar todo lo acabable y lo inacabable en siete meses, no sin ti no habría podido hacer ni la mitad de lo que hice. Gracias también por tu comprensión en esta última etapa en la que parece que nunca llega el final.

Cinco años dan para mucho, sobre todo cuando los pasas entre varios laboratorios y centros. Es mucha la gente que ha hecho que me alegre cada día de haber elegido el laboratorio donde estoy para hacer la tesis. En primer lugar todas y cada una de las personas que forman o han formado parte de mi grupo. Mil gracias a Susana, la Cañón, con la que he compartido todo este tiempo desde que empecé la tesis, y que no sólo es una gran compañera sino que a lo largo de estos años se ha convertido en una buena amiga. Muchísimas gracias por tu apoyo tanto científico como personal en tantos momentos, sobre todo en los más difíciles. Gracias por estar siempre dispuesta a ayudar en lo que sea con una sonrisa y por hacer que el día a día del laboratorio haya sido una alegría durante todo este tiempo. Muchas gracias a Miguel, con el que tanto he compartido. Muchas gracias por iniciarme en la microinyección y los clonajes, ese mundo oscuro y para-normal que nos ha traído de cabeza tantas veces. Muchas gracias por todos los buenos momentos pasados dentro y fuera del laboratorio y por tu apoyo, especialmente cuando estaba lejos de casa. Muchas gracias a Eva, gracias por tu empatía y tu apoyo incondicional en infinidad de ocasiones. Muchas gracias por dar alegría al trabajo diario,

por todo lo que aportas al grupo, no sólo a nivel de trabajo, que es mucho, sino a nivel personal. Es un placer trabajar contigo. Muchas gracias a Bea, con la que al final he coincidido en la poyata menos de lo esperable por estancias varias tuyas y mías. Gracias por tu apoyo, sobre todo en esta última etapa en la que parece difícil hacerse a la idea de que se acerca el final. Muchas gracias a Teresa, con la que únicamente he compartido un año, pero que parece mucho más. Muchas gracias por tu ayuda, por tu entusiasmo y tus ganas de trabajar. Gracias por dar alegría a este año y por la inyección de energía que me has dado cuando había poco de donde sacar. Tanto a Bea como Teresa, os deseo todo lo mejor para el resto de vuestra tesis, espero que terminéis con tan buen sabor de boca como yo lo hago. Gracias a Cristina por su apoyo en el tiempo que hemos coincidido en el labo. Muchas gracias también a aquellos que formaron parte de mis inicios en el Equipamiento 4 del IIB, Carlos y Cris Herranz, y también a Diego, que estuvo apenas un par de meses pero dejó huella. He tenido la suerte de dar con los mejores compañeros que podría desear, que han hecho que mi estancia en el laboratorio de M. Manzanera haya sido de las mejores etapas de mi vida. Muchas gracias a todos.

Casi un año de mi tesis lo pasé en el CSC de Londres, y la gente con la que compartí todo ese tiempo hizo que fuese una de las mejores experiencias que he vivido. Muchísimas gracias a toda la gente del laboratorio de Tristán, en especial a Aida, Margarida, Mel, Thom y Dan, por el cariño con el que me acogieron y por hacerme sentir en casa. Gracias especialmente a Mel, sin cuya ayuda en la disección de embriones no hubiera sido posible la realización de los arrays. Muchas gracias también por tu apoyo y disposición para prestar ayuda en todo momento. Muchas gracias a Margarida; gracias por tu amistad, por integrarme en la “colonia portuguesa”, por tu ayuda en tantas ocasiones y por todos los buenos ratos compartidos en el labo y fuera de él. Muchas gracias también a toda la gente del CSC que me ayudó, en especial a Marion Leleu, a la que no conozco personalmente pero cuyo análisis inicial de los datos del array me libraron de un presumible dolor de cabeza. Gracias también a Tatyana, que me dio consejos muy útiles cuando me enfagué en la derivación de las células TS.

Muchas gracias a toda la gente con la que he compartido esta última etapa en el CNIC. Gracias a toda la gente del departamento de Desarrollo Cardiovascular (y del Blastocisto y la Pata), a los grupos de Miguel, Nadia y Juanjo. Gracias en el plano científico por las aportaciones que habéis hecho a mi trabajo, porque cualquier idea, cuanto más se discute con gente que sabe, pues mucho mejor. Muchas gracias a Alberto, Jesús, Clara, Cati-Ana, Claudio, Verónica, Silvia, Cristina Cravería, las Lauras (Padrón y Carramolino), Adrian, Dani, Juanma, Johana y Rocío. Gracias a todos y cada uno de vosotros por hacer que la tercera norte del cnic tenga tan buen ambiente y que trabajar allí sea mucho más agradable. Muchas gracias a las tres secretarías del departamento, Sandra, Beatriz y Teresa, por estar siempre dispuestas a ayudar, por vuestra eficacia y porque gracias a vosotras nuestra vida es mucho relajada. Muchas gracias también a la gente de transgénesis, Chema, David y Luis Miguel, con los que he compartido mañanas de microinyección durante el último año, gracias por ese transgénico que me dio mi última alegría cuando había perdido la esperanza en los elementos reguladores.

Gracias a toda la gente de los servicios comunes del IIB y el CNIC, en especial a los encargados de nuestros animales, a Ana en el IIB, que puso orden en el caos, y a Abel en el CNIC, que nos hace la vida muy cómoda.

Mención especial merecen también todas aquellas personas con las que compartí mis primeros años en el IIB, que recuerdo con mucho cariño. Muchas gracias Itzi, apasionada de la ciencia a la que se lo pusieron demasiado difícil para seguir, me encantó compartir contigo la primera etapa de mi tesis. Muchas gracias a Jose, capaz de sacar el payaso que todos llevamos dentro y con el que convivir en el día a día supuso una de las mayores alegrías durante la época de Biomédicas. Muchas gracias por estos cinco años de tonterías, pero también de amistad, ya sea en Madrid o de estancias varias. Agradecimiento especial merecen Laura y Alberto, que han hecho las veces de mi familia londinense y con los que he compartido la adaptación a las lluvias permanentes, los veranos inexistentes y las moquetas omnipresentes. Muchas gracias por acogerme en vuestra casa aun a riesgo de que me apalancase (se han dado casos), por vuestro apoyo y por vuestra amistad.

Si estos agradecimientos estuviesen escritos en orden cronológico el primero al que debería agradecer el haber llegado hasta aquí sería Manu. Gracias a ti terminé la carrera teniendo claro que quería trabajar en desarrollo y gracias a ti elegí un laboratorio de evo-devo para hacer la tesis (espero que un día me perdonen el haberme centrado más en la devo). Muchas gracias por todos estos años de amistad, y los que nos quedan, aunque coincidamos a duras penas en la misma ciudad, y si me apuras en el mismo país o continente. Muchas gracias por tu apoyo científico y sobre todo personal, créeme que sin ti no estaría aquí. Muchas gracias también a Nacho (Maeso o google), compañero inseparable de Manu y que ha acabado siendo un buen amigo, además de inagotable fuente de conocimiento.

Muchas gracias a mis amigas de toda la vida, esas que nunca fallan sea lo que sea lo que estés haciendo, en este caso una tesis. Gracias a Ana, Andrea, Anna, Carla, Marta, Rebe, Aran y Mar. Gracias por años de amistad y por vuestro apoyo y comprensión siempre que lo he necesitado. En especial muchas gracias a Marta y Aran, siempre atentas, siempre dispuestas a ayudar y a dar apoyo pase lo que pase, ya estemos cerca o lejos. El trabajo de laboratorio puede llegar a ser muy frustrante, y es importante tener a gente cerca que te levante el ánimo. Para eso, para lo bueno y para lo malo, tengo unos amigos que me dan mucho más de lo que podría desear.

Dejo para el final de los agradecimientos a aquellos que han vivido de forma más cercana esta tesis o gran parte de ella y que en realidad representan el mejor apoyo que he tenido. Muchísimas gracias a mi familia: a mis padres, a mi hermano y a Eva. Tengo tanto por lo que daros las gracias que no sabría por dónde empezar. Muchas gracias por vuestro esfuerzo en intentar entender lo que hago y sobre todo por comprender lo importante que es para mí. Muchas gracias porque con vosotros a mi lado todo es posible.

A toda la gente que de una manera u otra forma parte de mi vida y de mi trabajo.

MUCHISIMAS GRACIAS

Bárbara

